

IJCSIS Vol. 13 No. 12, December 2015
ISSN 1947-5500

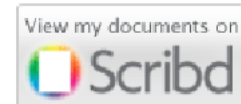
International Journal of Computer Science & Information Security

© IJCSIS PUBLICATION 2015
Pennsylvania, USA



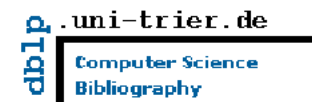
Cogprints

Google scholar



SciRate.com

CiteSeer^x beta



DOAJ DIRECTORY OF OPEN ACCESS JOURNALS



ProQuest

IJCSIS

ISSN (online): 1947-5500

Please consider to contribute to and/or forward to the appropriate groups the following opportunity to submit and publish original scientific results.

CALL FOR PAPERS

International Journal of Computer Science and Information Security (IJCSIS) January-December 2016 Issues

The topics suggested by this issue can be discussed in term of concepts, surveys, state of the art, research, standards, implementations, running experiments, applications, and industrial case studies. Authors are invited to submit complete unpublished papers, which are not under review in any other conference or journal in the following, but not limited to, topic areas.

See authors guide for manuscript preparation and submission guidelines.

Indexed by Google Scholar, DBLP, CiteSeerX, Directory for Open Access Journal (DOAJ), Bielefeld Academic Search Engine (BASE), SCIRUS, Scopus Database, Cornell University Library, ScientificCommons, ProQuest, EBSCO and more.

Deadline: see web site

Notification: see web site

Revision: see web site

Publication: see web site

Context-aware systems
Networking technologies
Security in network, systems, and applications
Evolutionary computation
Industrial systems
Evolutionary computation
Autonomic and autonomous systems
Bio-technologies
Knowledge data systems
Mobile and distance education
Intelligent techniques, logics and systems
Knowledge processing
Information technologies
Internet and web technologies
Digital information processing
Cognitive science and knowledge

Agent-based systems
Mobility and multimedia systems
Systems performance
Networking and telecommunications
Software development and deployment
Knowledge virtualization
Systems and networks on the chip
Knowledge for global defense
Information Systems [IS]
IPv6 Today - Technology and deployment
Modeling
Software Engineering
Optimization
Complexity
Natural Language Processing
Speech Synthesis
Data Mining

For more topics, please see web site <https://sites.google.com/site/ijcsis/>

arXiv.org

Google scholar

SCIRUS
search engine for science

ScientificCommons

Scribd

.docstoc
find and share professional documents

BASE
Bielefeld Academic Search Engine

CiteSeer^x beta

dblp.uni-trier.de
Computer Science
Bibliography

DOAJ
DIRECTORY OF
OPEN ACCESS
JOURNALS



ProQuest

For more information, please visit the journal website (<https://sites.google.com/site/ijcsis/>)

Editorial Message from Managing Editor

*The **International Journal of Computer Science and Information Security (IJCSIS)** is a peer reviewed journal publishing high-impact research which contributes new results and theoretical ideas in Computer Science & Information Security. The editorial board is pleased to present the December 2015 issue. The main objective is to disseminate new knowledge and latest research for the benefit of all, ranging from academia and professional communities to industry professionals. It especially provides a platform for high-caliber researchers, practitioners and PhD/Doctoral graduates to publish completed work and latest development in active research areas. We are glad to see variety of articles focusing on the major topics of innovation and computer science; IT security, Mobile computing, Cryptography, Software engineering, Wireless sensor networks etc. This scholarly resource endeavors to provide international audiences with the highest quality research and adopting it as a critical source of reference.*

IJCSIS archives all publications in major academic/scientific databases and is indexed by the following International agencies and institutions: Google Scholar, CiteSeerX, Cornell's University Library, Ei Compendex, ISI Scopus, DBLP, DOAJ, ProQuest, Thomson Reuters, ArXiv, ResearchGate, Academia.edu and EBSCO among others.

We thank and congratulate the IJCSIS team of editorial staff members, associate editors, and reviewers for their dedicated services to review and recommend high quality papers for publication. In particular, we would like to thank distinguished authors for submitting their papers to IJCSIS and researchers for continued support by citing papers published in IJCSIS. Without their sustained and unselfish commitments, IJCSIS would not have achieved its current premier status.

"We support researchers to succeed by providing high visibility & impact value, prestige and excellence in research publication."

For further questions or other suggestions please do not hesitate to contact us at ijcsiseditor@gmail.com.

*A complete list of journals can be found at:
<http://sites.google.com/site/ijcsis/>*

IJCSIS Vol. 13, No. 12, December 2015 Edition

ISSN 1947-5500 © IJCSIS, USA.

Journal Indexed by (among others):





Bibliographic Information

ISSN: 1947-5500

Monthly publication (Regular Special Issues)
Commenced Publication since May 2009

Editorial / Paper Submissions:

IJCSIS Managing Editor

ijcsiseditor@gmail.com

Pennsylvania, USA

Tel: +1 412 390 5159

IJCSIS EDITORIAL BOARD

| Editorial Board Members | Guest Editors / Associate Editors |
|---|--|
| Dr. Shimon K. Modi [Profile] Director of Research BSPA Labs, Purdue University, USA | Dr Riktesh Srivastava [Profile] Associate Professor, Information Systems, Skyline University College, Sharjah, PO 1797, UAE |
| Professor Ying Yang , PhD. [Profile] Computer Science Department, Yale University, USA | Dr. Jianguo Ding [Profile] Norwegian University of Science and Technology (NTNU), Norway |
| Professor Hamid Reza Naji , PhD. [Profile] Department of Computer Enigneering, Shahid Beheshti University, Tehran, Iran | Dr. Naseer Alquraishi [Profile] University of Wasit, Iraq |
| Professor Yong Li , PhD. [Profile] School of Electronic and Information Engineering, Beijing Jiaotong University, P. R. China | Dr. Kai Cong [Profile] Intel Corporation, & Computer Science Department, Portland State University, USA |
| Professor Mokhtar Beldjehem , PhD. [Profile] Sainte-Anne University, Halifax, NS, Canada | Dr. Omar A. Alzubi [Profile] Prince Abdullah Bin Ghazi Faculty of Information Technology Al-Balqa Applied University (BAU), Jordan |
| Professor Yousef Farhaoui , PhD. Department of Computer Science, Moulay Ismail University, Morocco | Dr. Jorge A. Ruiz-Vanoye [Profile] Universidad Autónoma del Estado de Morelos, Mexico |
| Dr. Alex Pappachen James [Profile] Queensland Micro-nanotechnology center, Griffith University, Australia | Prof. Ning Xu , Wuhan University of Technology, China |
| Professor Sanjay Jasola [Profile] Dean, School of Information and Communication Technology, Gautam Buddha University | Dr . Bilal Alatas [Profile] Department of Software Engineering, Firat University, Turkey |
| Dr. Siddhivinayak Kulkarni [Profile] University of Ballarat, Ballarat, Victoria, Australia | Dr. Ioannis V. Koskosas , University of Western Macedonia, Greece |
| Dr. Reza Ebrahimi Atani [Profile] University of Guilan, Iran | Dr Venu Kuthadi [Profile] University of Johannesburg, Johannesburg, RSA |
| Dr. Umar Ruhi [Profile] University of Ottawa, Canada | Dr. Zhihan Iv [Profile] Chinese Academy of Science, China |
| Dr. Shimon K. Modi [Profile] Director of Research BSPA Labs, Purdue University, USA | Dr Riktesh Srivastava [Profile] Associate Professor, Information Systems, Skyline University College, Sharjah, PO 1797, UAE |
| Dr. Jiliang Zhang [Profile] Northeastern University, China | |
| Dr. Vahid Esmaeelzadeh [Profile] Iran University of Science and Technology | |

TABLE OF CONTENTS

1. Paper 30111508: Student Performance Analysis using Large-Scale Machine Learning (pp. 1-6)

*Abdelmajid Chaffai, Abdeljalil El Abdouli, Houda Anoun, Larbi Hassouni, Khalid Rifi
RITM LAB ESTC , CED ENSEM, Hassan II University Casablanca, Morocco*

Abstract — The massive data generated by information systems in educational institutes need a special analysis in terms of efficient storage management and fast processing analytics, in order to gain deeper insights from the available data and extract useful knowledge to support decision making and improve the education service. In this work, we propose an intelligent system capable to predict the first year performance in departments chosen by the new enrolled student and recommend the suitable department. The system proposed will be helpful for educational decision-makers to reduce the failure rate by orienting the new students in department where they can succeed, and improve the enrollment management. Our approach is based on applying the machine learning techniques based on Map Reduce distributed computing model and the HDFS storage model. Decision tree and Support Vector Machines are used to build prediction model and clustering is used to develop a recommender system. To conceive our system, we relied on a real case study which is the results of students enrolled in the High School of Technology of Casablanca (ESTC), and collected between 2005 and 2014. The dataset is composed of two large data sources: the pre-higher education dataset and the result dataset of the first study year in the high school. Through experiments, the results generated by our system are very promising and the machine learning infrastructure implemented can be used for the future analytics on large variety of data sources.

Keywords: Machine learning; MapReduce; HDFS; Decision tree; Support Vector Machines; Prediction; Clustering; Recommender system.

2. Paper 30111513: Leveraging DEMF to Ensure and Represent 5ws&1h in Digital Forensic Domain (pp. 7-10)

*Jasmin Cosic, ICT Section of Police Administration, Ministry of the Interior of Una-Sana canton Bihac, Bosnia and Herzegovina
Miroslav Baca, Faculty of Organization and Informatics, University of Zagreb, Varazdin, Croatia*

Abstract — In this paper authors will discuss about metadata of digital evidence, digital chain of custody and framework for digital evidence management. A concretization of DEMF – Digital Evidence Management Framework will be presented. DEMF is a framework that was developed on conceptual – ontology level, and now programmed in java. One case study from real criminal case through DEMF will be presented.

Keywords – Digital evidence management framework, DEMF, Intererability, Digital evidence, Chain of custody

3. Paper 30111525: GCTrie for Efficient Querying in Cheminformatics (pp. 11-16)

*Yu Wai Hlaing, University of Computer Studies, Yangon, Myanmar
Kyaw May Oo, Faculty of Computing, University of Information Technology, Yangon, Myanmar*

Abstract — The field of graph indexing and query processing has received a lot of attention due to the constantly increasing usage of graph data structures for representing data in different application domains. To support efficient querying technique is a key issue in all graph based application. In this paper, we propose an index trie structure (GCTrie) that is constructed with our proposed graph representative structure called graph code. Our proposed GCTrie can support all types of graph query. In this paper, we focus on index construction, subgraph query and

supergraph query processing. The experimental results and comparisons offer a positive response to the proposed approach.

Keywords - Graph indexing and querying; Graph representative structure; Index; Subgraph query; Supergraph query

4. Paper 30111532: An Efficient Migration Scheduling Algorithm Based on Energy Consumption and Deadline in Cloud Computing (pp. 17-23)

*Sogol Davati, Department of Computer Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran
Ahmad Khademzadeh, Education and National International Scientific Cooperation Department, Iran Telecommunication Research Center (ITRC), Tehran, Iran*

Abstract — Cloud is one of the most important shareable environments which several customers are connected to it in order to access services and products. In this environment all the resources are available in an integrated environment and several users can simultaneously send their requests. In this condition some methods are required to schedule in an efficient way. Important condition is when a cloud server is overloaded or request of user has deadline constraint. In these situations to provide efficient service, migration occurred. Migration of virtual machines can be used as a tool for achieving balance. In other words when load of virtual machines in a physical machine increased, some of the virtual machines migrate to another physical machine so that virtual machines do not face any shortage of capacity. Virtual machine migration is used for many reasons like reducing power consumption, load balancing in virtual machines, online maintenance and capability of responding to requests with deadline constraint. In this thesis, we present an efficient algorithm for scheduling live migration of virtual machines which have applications with deadline constraint. This scheduling uses thresholds for migrating virtual machines in order to guarantee service level agreement (SLA). Simulations show that the proposed method acts better in number of received requests with success and level of energy consumption.

Keywords- Cloud computing; Virtual machine migration; Threshold; Cloudsim; SLA; Scheduling

5. Paper 30111533: EEG Discrimination of Rats under Different Architectural Environments using ANNs (pp. 24-31)

*Mohamed.I.El-Gohary, Department of Physics, Al-Azhar University, Cairo, Egypt
Tamer.A.A.Al-Zohairy, Department of Computer science, Al-Azhar University, Cairo, Egypt
Amir. M.M. Eissa, Department of Physics ,Al-Azhar University, Cairo, Egypt
Sally .M. Eldeghaidy, Department of Physics ,Suez canal University, Cairo, Egypt
Hussein.M.H.Abd El-Hafez, Department of Physics , Al-Azhar University, Cairo, Egypt*

Abstract — The present work introduces a new method for discriminating electroencephalogram (EEG) power spectra of rat's brain housed in different architectural shapes. The ability of neural networks for discrimination is used to describe the effect of different environments on brain activity. In this research the rats were divided into four groups according to the type of environmental shapes as: control (normal cage), pyramidal, inverted pyramidal and circular. The brain activities (EEG) were recorded from rats of each group. Fast Fourier Transform (FFT) analysis of EEG signals was carried out to obtain power spectra. Two different neural networks are used as classifiers for power spectra of the different 4 groups: multi-layer perceptron (MLP) with backpropagation and radial basis function RBF networks with unsupervised K means clustering algorithm. Experimental studies have shown that the proposed algorithms give good results when applied and tested on the four groups. The multilayer with backpropagation and radial basis function networks achieved a performance rate reaching 94.4 % and 96.67% respectively.

Keywords: EEG, Architectural shape, Artificial neural networks, Power spectrum, Backpropagation Algorithm, Radial basis function network.

6. Paper 30111534: Forward Error Correction for Storage Media: An Overview (pp. 32-40)

Kilavo Hassan, Nelson Mandela African Institution of Science and Technology, School of Computational and Communication Science and Engineering, P.O. Box 447 Arusha, Tanzania

Kisangiri Michael, Nelson Mandela African Institution of Science and Technology, School of Computational and Communication Science and Engineering, P.O. Box 447 Arusha, Tanzania

Salehe I. Mrutu, The University of Dodoma, College of Informatics and Virtual Education, P.O. Box 490 Dodoma

Abstract — As the adoption of Information and Communication Technology (ICT) tools in production and service rendering sectors increases, the demand for digital data storage with large storage capacity also increases. Higher storage media systems reliability and fault tolerance are among the key factors that the existing systems sometimes fail to meet and therefore, resulting into data loss. Forward error correction is one of the techniques applied to reduce the impact of data loss problem in digital data storage. This paper presents a survey conducted in different digital data storage companies in Dar es Salam, Tanzania. Data were collected and analyzed using Statistical Package for Social Sciences (SPSS). Secondary data were captured from user and manufacturer technical reports. It was revealed that data loss is still a predominant challenge in the digital data storage industry. Therefore, the study proposes the new storage media FEC model using locked convolutional encoder with the enhanced NTC Viterbi decoder.

Index Terms—Storage Media, FEC, NTC, Viterbi, RS

7. Paper 30111535: Certificate Based Hybrid Authentication for Bring Your Own Device (BYOD) in Wi-Fi enabled Environment (pp. 41-47)

Upasana Raj, Information Security and Cyber Forensics, Department of Information Technology, SRM University, Chennai, India

Monica Catherine S, Information Security and Cyber Forensics, Department of Information Technology, SRM University, Chennai, India

Abstract — Approval of the strategy, ‘Consumerization of IT’ by the organizations, does not only save money and increase business agility, but also improves employee productivity and satisfaction, lowers IT procurement, support costs and improves collaboration. Organizations have started to develop “Bring Your Own Device” (BYOD) policies to allow their employees to use their owned devices in the workplace. It’s a hard trend that will not only continue but will accelerate in the coming years. In this paper we focus on the potential attacks that can strike on BYOD when authenticating to a Wi-Fi network. It also enumerates the authentication protocols and methods. A proposal for stringent, indigenous hybrid authentication for a device that embraces BYOD strategy in a Wi-Fi enabled environment is proposed to the end of the paper.

Keywords — Bring Your Own Device (BYOD), Wi-Fi, Authentication, Authorization, Certificate.

8. Paper 30111536: Impact of Spectrum Sensing and Primary User Activity on MAC Delay in Cognitive Radio Networks (pp. 48-52)

Elham Shahab, Department of Information Technology Engineering, Nooretouba Virtual University of Tehran, Tehran, Iran

Abstract — Cognitive radio (CR) technology is used in wireless networks with the aim of handling the wireless spectrum scarcity. On the other hand, using CR in wireless networks has some new challenges and performance overheads that should be resolved. It is crucial to study the effects of the unique characteristics of cognitive radio over the various protocols. In this paper, a simulation-based performance evaluation is presented in the term of MAC delay in a CR user. The effects of spectrum sensing parameters (sensing duration and frequency), the primary users’ activity and the number of spectrum channels are investigated on the delay of MAC layer. The growth and decay rates of MAC delay are studied more in detail through the various NS2-based simulations. The results give some fruitful insights for formulating the delay of MAC based on the CR unique parameters.

Keywords-Cognitive radio; MAC; Delay; Primary user; Spectrum sensing

9. Paper 30111542: A Survey of Receiver Designs for Cooperative Diversity In The Presence of Frequency Offset (pp. 53-58)

*Sylvia Ong Ai Ling, Hushairi Zen, Al-Khalid b Hj Othman, Mahmood Adnan
Faculty of Engineering, University Malaysia Sarawak, Kota Samarahan, Kuching, Sarawak, Malaysia*

Abstract — Cooperative diversity is becoming a potential solution for future wireless communication networks due to its capability to form virtual antenna arrays for each node (i.e. user). In cooperative networks, the nodes are able to relay the information between the source and the desired destination. However, the performance of the networks (for instance – mobile networks, ad-hoc networks and vehicular networks) is generally affected by the mobility of the nodes. As the nodes' mobility rapidly increases, the networks are subjected to frequency offset and unknown channel properties of the communication links which degrades the system's performance. In a practical scenario, it is a challenging task and impractical for the relay and destination to estimate the frequency offset and channel coefficient especially in time varying environment. In this manuscript, a comprehensive survey of existing literature for receiver designs based on Double Differential (DD) transmission and Multiple Symbol Detection (MSD) approach is presented to eliminate the complex channel and frequency offset estimation.

Index Terms — *Cooperative Diversity, Double Differential, Frequency Offset, Multiple Symbol Differential Detection.*

10. Paper 30111544: Measuring and Predicting the Impacts of Business Rules Using Fuzzy Logic (pp. 59-64)

*Mohamed Osman Hegazi
Department of Computer Science, College of Computer Engineering and Science, Prince Sattam University, Saudi Arabia
Department of Computer Science, College of Computer Science and IT, Alzaiem Alazhari University Sudan*

Abstract - The relations between the constituent elements of the predominant activities may cannot be measured using the accurate measurement, the reason may be on the conflict or the effect of such rules with other factor, accordingly uncertainty measurement is needed. This paper presents a measurement and predictive fuzzy model that can be fit to work on such rules. The proposed model measures and predicting the impacts of such rules using fuzzy logic, it designed based on the concept of fuzzy logic, fuzzy set and the production rule. The model transforms if-then business rule to weighted fuzzy production rule, and then used this production rule for predicting and measuring the impact of the business rule. The model is tested using real data and provide considerable results.

Keyword: Fuzzy logic, Fuzzy set, Production rule, Business rule

11. Paper 30111549: Blind User Visualization and Interaction in Touch Screen: A Designer Perspective (pp. 65-72)

Abdelrahman H. Hussein, College of Computer and Engineering Science, University of Hail, Hail, Saudi Arabia

Abstract — In this paper, we describe how blind students views external system using an image map as a case study. We proposed two interaction techniques which allow blind students to discover different parts of the system by interacting with a touch screen interface. An evaluation of our techniques reveals that 1) building an internal visualization, interaction technique and metadata of the external structure plays a vital role 2) blind students prefer the system to be designed based upon their behavioural model to easily access and build the visualization on their own and 3) to be an exact replica of visualization, the metadata of the internal visualization is to be provided either through audio cue or domain expert (educator). Participants who used touch screen are novice users, but they have enough experience on desktop computers using screen readers. The implications of this study to answer the research questions are discussed.

Keywords- Blind; Visualization; Touch Screen; Accessibility; Usability; Image Map.

12. Paper 30111554: An Implementation of Android-based Steganography Application (pp. 73-76)

*Sultan Zavrak, Dept. of Computer Engineering, Duzce University, Duzce, Turkey
Seyhmus Yilmaz, Dept. of Computer Engineering, Duzce University, Duzce, Turkey
Huseyin Bodur, Dept. of Computer Engineering, Duzce University, Duzce, Turkey*

Abstract — The rapid development of smart phone technology has led to cheapening of the phone equipped with many advanced features such as sensors. One of the most widely used sensor in the phone is its camera. Although the photographs captured by camera can be shared via many ways, one of the most commonly used sharing methods is Multimedia Message Service (MMS) which allows transmission of files such as photographs, audio and video. A major disadvantage of MMS is that it does not provide sufficient safety mechanism and because of this, the data of the people who wants to hide confidential information from state-controlled systems can be easily monitored. In this study, unlike cryptography-based, a steganography-based mobile application that can embed the confidential information into an image, send it to receiver, and extract the confidential information from the image in the receiver side is developed. Besides, the performance data such as the embedding and extraction time of confidential information and experimental results of application are given.

Keywords— Multimedia Message Service, MMS, Steganography, Smart Phone.

13. Paper 30111558: Survey of the Adaptive QoS-aware Discovery Approaches for SOA (pp. 77-81)

Monika Sikri, Cisco Systems India Pvt. Ltd. SEZ, Embassy Tech Village, Panathur Devarabeesanahalli, Bangalore East Taluk Bangalore India

Abstract — Service Oriented Architecture is very commonly used as an architectural paradigm to model distributed integration needs. It provide the means of achieving organizational agility by building applications, which adapt to the dynamic needs of the business. Agility to Adapt is one of the key drivers for its growth. Its widespread adoption has led to the proliferation of multiple services offering similar functionality but different Quality of Service on the enterprise network. In real-time enterprise environment services are added and removed from the network on the fly. The service discovery approach does not only need to consider the QoS of other similar services but also the impact of dynamic and unpredictable system behavior on QoS. In view of this, there is a need for adaptive discovery approach that can accommodate these run-time changes and keep the system functioning despite the unpredictability. As part of this work we have reviewed existing works in Adaptive QoS-aware discovery for SOA systems to understand the gaps and need for future research.

14. Paper 30111563: M-Learning for Blind Students Using Touch Screen Mobile Apps Case Study - Special Education in Hail (pp. 82-88)

*Abdelrahman H. Hussein (1), Majed M. AlHaisoni (1), Ashraf A. Bany Mohammed (2), Mohammed Fakrudeen (3)
(1) College of Computer Sciences and Engineering, University of Hail, Hail, Kingdom of Saudi Arabia
(2) School of Business, The University of Jordan, Amman, Jordan
(3) Dept. of Computing and Technology, Anglia Ruskin University, Chelmsford, United Kingdom*

Abstract - The relative newness of the touch-screen (TS) based device creates a phenomenon unique and unstudied in the academic environment with regard to blind students dependent on Braille. This qualitative research study explores how the use of a multi-modal touch-screen based device affects the academic environment for totally blind students using YouTube videos. The pilot program (android app) included a prototype for the English course offered to fifth grade level pupils attending primary school in Hail, KSA. Data collected from students through a survey and focus group interviews and from the faculty through individual interviews was coded and organized according to the research questions. Findings analysis was organized by way of the study's conceptual framework: (a) substitution of

Braille course materials with YouTube video lessons (b) accessibility and usability of the developed prototype. Findings concluded that the majority of students in this study perceived YouTube course materials on a touch-screen based device (using android app) to be as good as, or better, than Braille course materials, the multi-modal functionality of the touch-screen based device augmented personal study and classroom learning, and the personal use positively contributed to academic use of the device.

Keywords- Accessibility; Usability; Touch screen; M-learning; YouTube videos; and Blind students

15. Paper 30111506: A Fitness-Gossip Routing Protocol for Saving Energy in Wireless Sensor Networks (pp. 89-95)

Kareem Radi Hassan, Department Of Computer Science, University of Basrah, Basrah, Iraq

Abstract - Gossiping is traditional routing scheme commonly used in WSNs to transport data in a multi hop manner because of limited radio range and energy constraints. Gossiping protocol is simple and do not involve additional devices, but at the same time it faces several deficiencies when it used in WSNs. This paper describes an efficient technique to transmit the message to the sink node. The main idea behind the new protocol which is called (Fitness-Gossiping Protocol FGP) which is a modification of Gossiping protocol by using a fitness function to select the optimum next node, when the optimum next node is selected the data is transmitted to the next node. We discussed how the new approach saved the energy of the network and to achieve maximize the network lifetime in comparison with its counterparts. In the same time, the Fitness-Gossiping protocol provides the balanced energy between nodes.

Index Terms— Gossiping, Fitness- Gossip, Network lifetime, routing, Wireless sensor networks (WSNs).

16. Paper 30111528: Computing the Dynamic Reconfiguration Time in Component Based Real-Time Software (pp. 96-104)

*Asghar Farhadi, Computer depart, Islamic azad university Arak branch, Arak, Iran
Mehran sharafi, Computer depart, Islamic azad university Najafabad Branch, Najafabad ,Iran*

Abstract - New microcontrollers with enhanced application capabilities for more complex scenarios are developed. However, the settings are volatile and ever-changing environment requires permanent systems that compatible with the new conditions. Dynamic re-configuration has a powerful mechanism for the implementation of the adaptation strategy. Real-time control system is one of the challenges for the implementation of dynamic re-configuration software. In previous works the Adapt.NET is adapted in framework of the implementation of component-based applications. A new web-based test complete reconfiguration is proposed here with a limited time. The application dynamically re-configures the client-side component compatibility in case of failure of the component parts. In this article the timing behavior of the implemented dynamic reconfiguration algorithm is analyzed. The manner of the hybrid component-based applications adaptation during environmental condition changes are described as well. In this article ,in order to predict the implementation time the behavior of dynamic reconfiguration algorithm and the manner of real-time planning that can be adapted to environmental changes are assessed, as well as the correlation of reconfiguration during the deadline period.

Key words: Dynamic reconfiguration, Blackout, Reconfiguration time, Adaptation, state

17. Paper 30111556: Intrusion Detection Systems: A Novel Multi-Agent Specification Method (pp. 105-110)

*Eddahmani Said & Rahal ROMADI, University Mohammed V Rabat, Faculté des Sciences de Rabat, Morocco
Bouchaib BOUNABAT, University Mohammed V Rabat, ENSIAS, BP 713, Agdal Rabat, Morocco*

Abstract - Intrusion detection systems play a major role in information systems security. However, the complexity of a network flow can cause the outbreak of numerous false alarms, commonly called false positives or cannot detect malicious traffic, i.e. false negatives. Our objective in the present paper is to put forward a novel method that

combines several types of monitoring, intrusion and malware detection tools whose main role is to reduce the rate of false positives/negatives.

Keywords: malware detection system, Reactive Agent, specification, Formal methods.

18. Paper 30111517: A Brief Overview of Cooperative MAC Protocols in Wireless Networks (pp. 111-116)

Ferzia Firdousi, Department of Electrical Engineering, National University of Sciences and Technology, Rawalpindi, Pakistan

Abstract - Cooperative diversity is a fairly recent technique. It allows the benefits of space diversity to be incorporated into the wireless communications network. Whereas it has found widespread use in network, MAC and physical layers, this paper highlights its implementation in the MAC layer as a researchable area. A wide number of techniques have been deployed in the MAC layer with a scope for good cooperation for the modern wireless networks. While all of them have been found to be useful, a few of them have disadvantages as well. In this paper, we present a brief overview of the MAC protocols deployed in the cooperative diversity of wireless communication. This survey is compact enough to understand properly but detailed enough to provoke research ideas.

Keywords - Cooperative Diversity, Mac Protocol, Energy Consumption

19. Paper 30111555: Architectural Framework for Inter-Agency Intelligence Information Sharing Knowledge Based System in Botswana (pp. 117-131)

*Ezekiel U Okike, Computer Science Department, University of Botswana, Gaborone, 00000, Botswana
T. Leburu-Dingalo, Computer Science Department, University of Botswana, Gaborone, 00000, Botswana*

Abstract - The 21st century is witnessing increasing wave of terrorist and crime activities. This has made the need for intelligence-led policing the focus of most national governments including Botswana. The main objective of this paper is to propose an architectural model for intelligence-Led policing in Botswana which will provide support for all intelligent agencies in the country namely the Botswana Police Service (BPS), Directorate of Intelligence and Security (DISS), Directorate of Corruption and Economic crime (DCEC), and Criminal Investigation Department (CID). The model provides for inter agency information sharing using appropriate modern technologies in order to enable the agencies access to useful crime data which will enhance their ability to make informed decisions and take appropriate actions against cyber and economic crimes.

Keywords: Cyber and economic crime, Security, intelligence policing, architectural framework, Knowledge base, Information sharing, system model.

20. Paper 30111518: A Comprehensive Review of Different types of Cooperative MAC Layer Protocols (pp. 132-141)

Ferzia Firdousi, Department of Electrical Engineering, National University of Sciences and Technology, Pakistan

Abstract - For physical layer, cooperative systems have been devised which aim at increasing diversity gain achieved by spatial diversity. This advantage can be mapped onto the MAC layer to achieve throughput increment, faster transmission rate, reduction in power and energy consumption, and a large network coverage area. However in the race to achieve a good MAC layer protocol, many new problems are created, including a redundant use of relay nodes, increase in energy wastage and increase in delay etc. In order to understand the true capabilities of cooperative communication at the MAC layer, many protocols need to be studied and their drawbacks identified, so that the upcoming research can be improved. This paper clinches research regarding these different types of cooperative MAC protocols by summarizing and analyses them. The analytical result, discourses issues like which relay node is the optimal one, how can a system be made energy efficient and which methodology to be followed for

forwarding a data packet in case of transmission failure, and explains them in detail to allow room for future, new and improved research.

Keywords: Cooperative, MAC, Relay selection, Diversity gain, Energy efficiency

21. Paper 30111562: Using Cohesion and Capability Maturity Model Integration (CMMI) as Software Product and Process Quality criteria: A Case Study in Software Engineering Practice in Botswana (pp. 142-149)

Ezekiel U Okike & Motsomi Rapoo, Computer Science Department, University of Botswana, Gaborone, 00000, Botswana

Abstract - The need to ensure quality in software engineering practice necessitated the introduction of software measurement and other quality standards introduced by the software engineering institute. The appropriate introduction of quality standards however does not assume that practitioners easily adapt to its utilization without assistance by experienced professional. This study aims to investigate software engineering practice in Botswana in order to assist software companies understand and use software engineering quality measures and standards. 14 software companies were identified out of which 7 indicated interest in participating in this study. The result indicates that most of the participating companies are yet to satisfy the Capability Maturity Model Integration (CMMI) 18 key performance areas at the 5 levels. Of the 5 companies which indicated that they are using CMM/CMMI standard, only 1 company satisfies 100% CMMI requirements. The study reveals the need to launch a programme to bring CMMI into Botswana and to train software companies on the use of appropriate software metric tools in order to ensure software quality in Botswana.

Keywords: Software engineering, software measurement, software quality, cohesion, capability maturity model integration

22. Paper 30111553: Trust: Models and Architecture in Cloud Computing (pp. 150-153)

*Usvir Kaur, I. K. Gujral Punjab Technical University, Jalandhar
Dheerendra Singh, Dept. of Computer Science & Engineering, Shaheed Udham Singh College of Engineering and Technology, Tangori, Mohali, Punjab, India*

Abstract - In today era, Cloud Computing has become an emerging technology. Various service providers are available in market with different configurations. It becomes very difficult for users to choose between various cloud service providers. Trust plays a vital role in field of cloud computing services such that it enables users to choose particular cloud service provider from available service on basis of how this technology behaves in past. This paper discusses various parameters and trust model framework. Different trust models for single web service and various parameters used by them for calculating trust are reviewed.

Keywords: cloud computing, trust model architecture, security.

23. Paper 30111557: Towards a mobile payment market: A Comparative Analysis of Host Card Emulation and Secure Element (pp. 156-164)

*Pardis Pourghomi, Department of Computer Science, American University of the Middle East, Kuwait
Pierre E. Abi-Char, Department of Computer Science, American University of the Middle East, Kuwait
Gheorghita Ghinea, Department of Computer Science, Brunel University, Kingston Lane, Uxbridge, Middlesex, UB8 3PH, United Kingdom*

Abstract — The considerable existing potential for mobile payments adoption shows that businesses are interested to increase the number of electronic transactions while consumers are attracted to convenient ways for fast and accessible banking. Nevertheless, the belief that the value of the Near Field Communication technology has not yet

been fully recognized - particularly in the consumer marketplace - persists. However, the introduction of Android 4.4 operating system namely 'KitKat' has pushed the Near Field Communication (NFC) market towards Android devices with the recently proposed Host Card Emulation (HCE) technology. Moreover, there are various debates about the ways in which mobile payment processes should be managed. Currently, the most recognized and accepted methods for managing the mobile payment processes are the traditional Secure Element (SE) approach and the Host Card Emulation which has lately become a crucial topic for key industry players. This paper describes the aspects of moving forward with mobile wallets. More specifically, a broad discussion is developed to consider the pros and cons of both approaches. Correspondingly, a detailed analysis is carried out centred on the security and adoption issues that these approaches may raise.

Keywords - Near Field Communication; Secure Element; Host Card Emulation; Mobile transaction.

24. Paper 30111564: Performance Analysis of Sybil Decline: Attack Detection and Removal Mechanism in Social Network (pp. 165-171)

*Deepti S. Sharma, Department of CSE, LKCT, Indore, M.P, India
Dr. Sanjay Thakur, Department of CSE, LKCT, Indore, M.P, India*

Abstract - Peer to peer system involves communication between the two directly connected hosts. As it supports the open communication medium it suffers from various security threats from remote computing elements. To avoid these threats various systems and integrated solutions are proposed over the last few years. One of such well known problem is Sybil attacks in which user performs unauthorized or malicious activities by creating the fake identities over the webs. One way to keep such threat away from the system is to design a centralized trusted authority system. Among its protection approaches Sybil Defender, Sybil Limit, Sybil Guard and Sybil Shield are some well known tools. After analyzing these tools and their respective mechanism we have found that they somewhere lack the associated trust computation issues. For improving such issues we have suggested a novel approach named as Sybil Decline. This paper gives a performance evaluation of suggested approach on some performance monitoring parameters. Extensive analysis and number of experiments are performed here to prove the results authenticity and effectiveness of the suggested approach.

Keywords - Sybil, Security, Peer to peer system, trusted authority

25. Paper 30111561: Cost and Performance Based Comparative Study of Top Cloud Service Providers (pp. 172-177)

Irfan Bari (1), Salaser Babu (2), Muhammad Munwar Iqbal (3), Yasir Saleem (4), Zahid Ali Masood (1), (2), (3) Department of Computer Science, University of Engineering and Technology, Taxila (3), (4) Department of Computer Science & Engineering, University of Engineering and Technology, Lahore

Abstract — Recent boom in the cloud computing industry has caused a shift in the information technology industry and has affected the way how information and data is stored and shared among the enterprise. The advent of social applications also demands the availability of resources that can be shared among the others. Cloud based architecture has made it possible for enterprises to utilize the computation power that was not available in the past. This paper takes a look and compares the top available service providers on the basis of the cost for each computing model as well takes a look at the performance by measuring the response time. It is observed that at all these service providers and elaborated the comparison of all based on their available architectures.

Keywords — Comparison of Cloud Operators, Cloud Computing, Azure, RackSpace, Amazon Web Services

Student Performance Analysis using Large-Scale Machine Learning

Abdelmajid Chaffai, Abdeljalil El Abdouli, Houda Anoun, Larbi Hassouni, Khalid Rifi
RITM LAB ESTC
CED ENSEM Hassan II University
Casablanca, Morocco

Abstract— The massive data generated by information systems in educational institutes need a special analysis in terms of efficient storage management and fast processing analytics, in order to gain deeper insights from the available data and extract useful knowledge to support decision making and improve the education service. In this work, we propose an intelligent system capable to predict the first year performance in departments chosen by the new enrolled student and recommend the suitable department. The system proposed will be helpful for educational decision-makers to reduce the failure rate by orienting the new students in department where they can succeed, and improve the enrollment management. Our approach is based on applying the machine learning techniques based on Map Reduce distributed computing model and the HDFS storage model. Decision tree and Support Vector Machines are used to build prediction model and clustering is used to develop a recommender system. To conceive our system, we relied on a real case study which is the results of students enrolled in the High School of Technology of Casablanca (ESTC), and collected between 2005 and 2014. The dataset is composed of two large data sources: the pre-higher education dataset and the result dataset of the first study year in the high school. Through experiments, the results generated by our system are very promising and the machine learning infrastructure implemented can be used for the future analytics on large variety of data sources.

Keywords: machine learning; MapReduce; HDFS; Decision tree; Support Vector Machines; prediction; Clustering; recommender system

I. INTRODUCTION

Educational Data Mining (EDM) is defined on the site community www.educationaldatamining.org as: "Educational Data Mining is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in". Several methods and approaches are used in Educational data mining research including data mining, machine learning, and exploratory data analysis in order to understand data and model the hidden relationships among data sets. These methods include classification, clustering, and association rules.

Nowadays, computer systems in educational institutes are producing gradually year by year a huge quantity of data. There is a great interest to use these massive data to analyze,

understand and extract or bring out some solutions to problems that affect educational system such the failure in first year, and help the right decision making.

Managing and analyzing huge volume of data in the machine memory is not an option, it needs a special infrastructure. We implement a hybrid distributed solution for both storage and analysis on cluster, this solution will ensure a scalable storage of educational data and fast processing analytics.

In this work, we use many techniques of machine learning and exploratory data analysis to produce an interactive dashboard dedicated for decision-makers in the aim to summarize the main characteristics of student data with visual methods, predict the performance of the new enrolled students in the first year and recommend the department where they can have good results.

II. MACHINE LEARNING

Machine Learning is a field of study and development methods capable to learn from data and transform the learning into action to accomplish a defined task. A formal definition cited by Tom Mitchell [1] is "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E".

Machine learning is divided in three categories depending on type of data available and the desired task to accomplish.

- a. Supervised learning: input data in a learning algorithm is a set of examples or observations whose response or target is known. The principle of algorithm is to model and optimize the relationship between target attribute and other attributes from the examples of the set of training data, and the model is applied on new unlabeled data. Supervised machine learning is widely used in building the predictive models.
- b. Unsupervised learning: the target attribute is not specified in input data to give to the learning algorithm. Unsupervised learning is about analyzing the unlabeled data and finding the data structure. It is widely used in building the descriptive models such clustering, association rules, recommender system, and pattern discovery.

- C. Reinforcement learning: a machine accomplishes a task and interacts with dynamic environment by producing actions without a human intervention. Reinforcement learning is used in many applications such as autonomous car.

A. Classification with Decision tree

Decision tree learner use a method known as "divide and conquer" to build a model from a set of attributes in tree structure based on logical decisions. The built tree is composed of one root, branches, interior nodes and leaves nodes. Each interior node represents test made on one value of the input attributes, each branch is the result of decision's choice, and each leaf is a terminal node that represent target value or class label in classification. To build a decision tree, an appropriate test is applied to divide the samples into homogeneous subsets of similar classes. The perfect test would divide the data into segments of single class, they are considered pure.

There are many implementations of decision tree algorithms, the most recent and known is C5.0 [2] algorithm, this algorithm was developed by J.Ross Quinlan. C5.0 is a successor and improvement of his prior algorithm C4.5 [3].

B. Classification with Support Vector Machines

Based on the work of Corinna Cortes and Vladimir Vapnik in 1995 [4] Support Vector Machine (SVM) model is a representation of the examples as points in space. It uses boundaries called hyperplanes to separate a set of instances into homogeneous groups in d-dimensional space, the hyperplane is a line in 2D space but in multi-dimensional space it is a flat surface. SVM learner attempts to create the greatest separation to the nearest training-data points from each class by defining the Maximum Margin Hyperplane (MMH). The points of any class that are the closest to the MMH are called the support vectors. In real-world application, the classes are not linearly separable. SVM uses kernel functions and transform the data into a higher dimensional feature space, using this process a nonlinear relationship can be linear.

C. Clustering

Clustering is an unsupervised machine learning technique that divides the data in natural groups known as clusters of similar elements; it provides additional and new information about data.

The most often used algorithm is k-means whose principle is:

- a) Partition data in k groups into a predetermined number of clusters.
- b) Choose random k points as cluster centers.
- c) Assign examples to their nearest cluster center according to the Euclidean distance function.
- d) Update the centroids for the cluster by calculating the mean value of the points assigned to the cluster.
- e) Repeat phases 2, 3 and 4 until the same points are assigned to each cluster.

III. RELATED WORK

Brijesh Kumar Baradwaj and Saurabh Pal [5] used classification with Decision tree based on ID3 by selecting a sample of 50 Master of computer applications students with 7 predictors, to evaluate and predict the student's performance at the end of the semester; the knowledge is extracted from classification in the form of IF-THEN rules.

Zlatko J. Kovačić [6] conducted a study by selecting an enrollment dataset of 450 students to analyze the influence of socio-demographic variables (age, gender, ethnicity, education, work status, and disability) and study environment (course programme and course block) in predicting the student's success, using four different methods of classification with four different classification trees (CHAID, exhaustive CHAID, QUEST and CART) and logistic regression. Data analytics were performed with SPSS 17 and Statistica 8.

M.Sindhuja et al., [7] conducted a study by selecting a dataset collected from a tutoring system to perform analytics based on related attributes such as behavior, attitude and relationship. Prediction of good behavior, average attitude and good relationship with faculty members and tutors was fitted with hierarchical clustering analysis using DBScan clustering, data analytics were performed with WEKA tool.

Krina Parmar et al., [8] proposed a framework to predict student performance based on attributes like past results, mid semester results, academic punctuality and performance in various tests. Datasets are distributed over servers, metadata are managed by a middleware that allows applying the data mining algorithms on the datasets like classification, clustering, association rule. The final model is a merging of results of each local model performed where the data are located.

IV. TOOLS

1) Hadoop Framework

The Apache Hadoop [9] is an open-source software framework written in Java. It is a set of algorithms designed to make the storage and the processing distributed across clusters of commodity computers, using a specific model of computing called MapReduce, and a specific model for storage called HDFS. Hadoop is designed with a very high degree of fault tolerance, and resolve problems of dead nodes with replication, which is one of the most powerful characteristic in this framework comparing with other existing distributed systems.

1.1) HDFS: The Hadoop Distributed File System inspired from Google file system GFS is a scalable distributed file system for large distributed data-intensive applications [10]. It provides scalable and reliable data storage. HDFS is master/slave architecture, composed of two main components: the NameNode is the master and the DataNode is the slave. The NameNode contains the metadata of all files system including permissions, locations of data, and information of replicated blocks and manage the access to file by the client. It divides the input data into blocks respecting the block size in the configuration files (from 64Mo to 128Mo) and store data

blocks into DataNodes in an intelligent manner, in order to optimize the network bandwidth. The NameNode orchestrates the storage operation occurred in data nodes while the DataNode stores data and serves it to client requests.

1.2) MapReduce is a programming model and an associated implementation for processing and generating large data sets [11]. It's also master/slave architecture, it allows to write applications in a parallel manner over a large data sets across the cluster, the design of a distributed program is divided in two main phases Map and Reduce. Management of the distributed process is ensured by the Job Tracker daemon in the master node, and the execution of tasks map and reduce is ensured by the Task Trackers daemons in the slaves nodes.

2) Hadoop streaming

The Java MapReduce API is the standard option which allows writing MapReduce programs in java. To write MapReduce jobs in other programming or scripting languages, we use an utility called Hadoop Streaming [12]. It is a generic API that provides options which allows writing and running executable scripts using standard input and output over the Hadoop cluster. Hadoop streaming supports C++, PHP, Perl, Shell, Ruby, and R.

3) R with hadoop

Analyzing large dataset in a single machine with R [13] algorithms is constrained by the limitations of memory size. Combining R power in analytic with the power of Hadoop cluster is an efficient solution to extend R processing capabilities on massive data.

In this work, we use two packages rmr2 [14] and rhdfs [15] developed by Revolution Analytics to connect R environment with the Hadoop cluster. Rmr2 package allows writing MapReduce programs in R, and depends on the Hadoop streaming. rhdfs package provides HDFS file management in R. The mapreduce() is the main function in rmr2 package, its syntax with arguments is:

mapreduce(input, output, input.format, output.format, map, reduce).

The input and output arguments are the paths of input and output HDFS data. The input.format and output.format arguments are in string format with additional arguments to specify the format of the files. The map and reduce arguments are the two steps of MapReduce algorithms implemented in R language.

4) Hive

Hive is data warehouse software developed initially by Facebook [16], it allows users query and manage large datasets stored in Hadoop distributed cluster using a language called Hive Query Language (HQL) very similar to standard ANSI SQL. Hive provides an abstraction on top of Hadoop, it translates the ad-hoc queries in one or more MapReduce jobs,

which are then submitted to the Hadoop cluster for execution, and returns the results to the user. Hive uses metastore service to store metadata for tables in relational database. By default the metastore uses derby which is embedded Java relational database, it can't be used in multiuser environment, we use MySQL database as a metastore. Hive provides a service that allows any client to submit requests and retrieve results, this service is ensured by the hiveserver2 [17] based on Thrift RPC [18]. To connect R environment with Hive warehouse, we use the Rhive [19] library which allows to access, manage and process over Hive data warehouse in R. In this work, we use Hive to build tables from dataset stored in HDFS, and integrate visualization tools to perform the exploratory data analysis.

V. METHODOLOGY OF EXPERIMENT

1) Environment experiment

We deployed a Hadoop cluster on 11 machines with the following configuration: Intel(R) Core(TM) i5-3470 CPU 3.20GHz (4CPUs), 8GB Memory, 300GB hard disk. One machine is designated as master (Namenode and JobTracker), the others are workers (DataNodes and TaskTrackers). R environment, with rmr2 and rhdfs packages are installed on all nodes. Hive and Rhive package are deployed on master . The Platform architecture is shown in Figure 1:

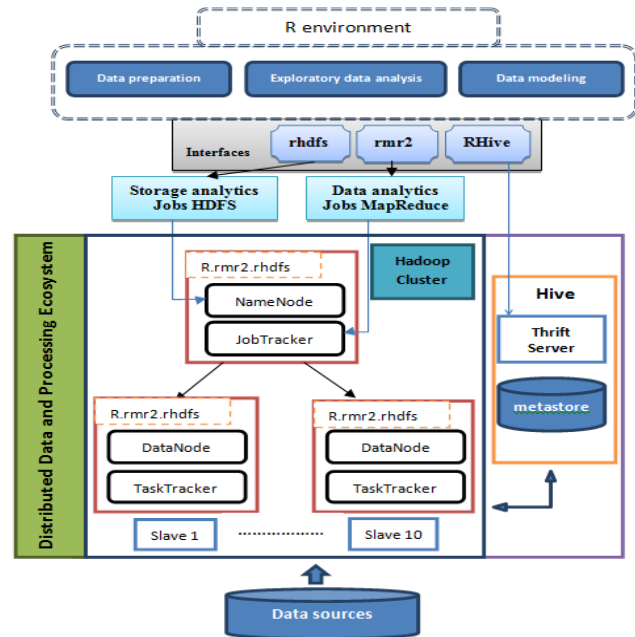


Figure 1. Distributed analytics Platform architecture for Machine learning using R environment

2) Problem definition and goals

The high school of technology of Casablanca (ESTC) has five departments:

- a) Computer Engineering (GI)
- b) Electrical Engineering (GE)
- c) Chemistry Process Engineering (GP)
- d) Mechanical Engineering (GM)
- e) Technical Management (TM)

A student who wants to enroll in the 1st year of the high school has to complete an on line form. This form proposes a list which can contain until three departments (among 5) that the student can join. The goal of our work is to guide the student, according to some results he obtained in the baccalauréat exam (baccalauréat or bac is a certificate /diploma awarded in General or Technical Secondary level in Moroccan education system), to choose the department where he can obtain better results. The admission is based on the overall average obtained in this exam.

The results we used are the marks obtained in mathematics, physics-chemistry and French language. We give interest to French because, in pre-higher education, the majority of courses are taught in Arabic which is the maternal language, while in higher education, all courses are given in French, so mastering French is necessary. The main goals of this data analytics problem are as follows:

- a) Predict the result that a new student could obtain in each department where he is allowed to enroll in.
- b) Recommend the most suitable department to the student.

3) Design data requirement and data collect

Our work uses two data sources which are related to the problem.

- a) A dataset containing the results obtained by students in baccalauréat exam of the last 10 years between 2005 and 2014. The dataset consists of a csv file whose headers are presented in table I:

TABLE I. ATTRIBUTES DESCRIPTION OF DATASET CORRESPONDING TO THE RESULTS IN BACCALAURÉAT

| Attribute | Description | Type (domain) |
|------------|---|-----------------------------------|
| id | Student's identity | Numeric |
| year | year of obtaining the baccalauréat | Integer |
| gender | Student's sex | Nominal(binary:Female, male) |
| academy | Origin region of the student | Nominal (with 16 values(regions)) |
| bac_type | Baccalauréat type | Nominal (with 5 values) |
| bacAverage | Overall average obtained in the baccalauréat exam | Continuous (from 0 to 20) |
| frMark | Mark in French language | Continuous (from 0 to 20) |
| mathMark | Mark in mathematics | Continuous (from 0 to 20) |
| pcMark | Mark in physics-chemistry | Continuous (from 0 to 20) |

- b) Dataset extracted from a database containing all the results obtained by the last 10 promotions in the 1st year of the high school (ESTC). The dataset consists of a csv file whose headers are described in Table II:

TABLE II. ATTRIBUTES DESCRIPTION OF DATASET CORRESPONDING TO THE RESULTS OF THE LAST 10 PROMOTIONS

| Attribute | description | Type(domain) |
|------------|-----------------------------------|---------------------------------------|
| id | student's identity | Numeric |
| year_study | year of study | Integer |
| dept | department | Nominal (with 5 values (departments)) |
| fyAverage | first year average of all results | Continuous (from 0 to 20) |

4) Preprocessing data

To process data, we begin by storing in the HDFS distributed storage system the two files corresponding to the two datasets presented above. Then, we write a R program using distributed mode and Rhadoop integration to merge the two csv files in one csv file named data2005_2014. The merge uses the student's id and is completed by the equijoin() function. The headers of the file data2005_2014.csv are: id, year, gender, academy, bac_type, bacAverage, frMark, mathMark, pcMark, dept, fyAverage.

To exploit the resulting file stored in HDFS, we begin by writing a MapReduce program in R by using make.input.format() function which takes the path of data2005_2014.csv as input, and translates this csv file into format that can be read by rmr2 functions. Then, we start the different data analytics across the Hadoop cluster.

5) Training models on data

A. Classification

To train the classification models we discretize fyAverage attribute which is continuous by converting it into categorical attribute according to the following partitioning:

- Risk : if $fyAverage < 12$ (to succeed, a student must get a global average greater or equal to 12)
- Acceptable : if $12 \leq fyAverage < 14$
- Good: if $14 \leq fyAverage$

fyAverage is defined as response variable.

1) Decision tree Model

First, from data2004_2015 we create a randomly ordered dataset, then we divide it into two portions. The first portion, containing 90 percent, is used as training dataset to build the classification model. The second, containing 10 percent, is used as test dataset to evaluate the performance of the model. The C50 package [20] contains an implementation of C5.0 decision tree algorithm, which uses entropy for measuring the impurity of a set of samples according to their target classification. The entropy of a set S of samples is:

$$Entropy(S) = \sum_{i=1}^C - p_i \log_2(p_i)$$

Where c is the number of different class levels, and p_i is the proportion of samples of class level i in S . The information gain of an attribute A is the reduction in entropy, which can be expected if a split is made on the basis of this attribute. It is calculated as the difference between the entropy before the split $S1$ and the partitions resulting from the split $S2$

$$InfoGain(A) = Entropy(S1) - Entropy(S2)$$

By maximizing the information gain and minimizing the entropy, the attribute with the highest information gain is selected to create homogeneous groups.

We use the C5.0() function in mapreduce functions with default settings. Then the model resulted is applied to the test dataset by using the predict() function. Table III shows its evaluation performance.

TABLE III. PERFORMANCE OF C5.0 MODEL WITH DEFAULT SETTINGS.

| Model | accuracy | Kappa statistic | F-measure good | F-measure acceptable | F-measure risk |
|-------|----------|-----------------|----------------|----------------------|----------------|
| C5.0 | 89.3% | 0.84 | 0.92 | 0.82 | 0.92 |

Adaptive boosting is one of features of C5.0 based on the work of Rob Schapire and Yoav Freund [21]. This feature allows to control the number of the boosting iterations. By default C5.0 starts with trials parameter equal to 1 iteration. Adding 10 trials which imply building 10 separate decision trees instead of one, we improved accuracy, from 89.3 to 97%. The boosted model performance is shown in Table IV.

TABLE IV. PERFORMANCE OF BOOSTED C5.0 MODEL

| Model | accuracy | Kappa statistic | F-measure good | F-measure acceptable | F-measure risk |
|-------|----------|-----------------|----------------|----------------------|----------------|
| C5.0 | 97% | 0.94 | 0.97 | 0.90 | 0.97 |

2) SVM model

SVM learner requires each data sample is described with numeric features. We use 1-of-K coding to convert all categorical attributes into numeric data. To produce the training data set and the test dataset, we use the same method in previous classification with Decision tree.

kernlab package [22] is used to build different SVM models, by changing the value of kernel argument in ksvm() function. We start from simple linear kernel "vanilladot", to complex non-linear kernels like "rbfdot" Radial Basis Function kernel (Gaussian kernel), "polydot" Polynomial kernel, "tanhdot" Hyperbolic Tangent kernel, "laplacedot" Laplacian kernel, and "besseldot" Bessel kernel.

The model with Radial Basis Function kernel presents a good performance comparing to others kernels as shown in Table V.

TABLE V. PERFORMANCE OF SVM MODEL WITH RADIAL BASIS FUNCTION KERNEL

| Model | accuracy | Kappa statistic | F-measure good | F-measure average | F-measure risk |
|-------|----------|-----------------|----------------|-------------------|----------------|
| SVM | 80% | 0.69 | 0.90 | 0.63 | 0.85 |

3) Conclusion

Both tables IV and V show clearly that the model C5.0 allows to obtain better results than SVM model. It presents good prediction performances with accuracy equal to 97%. To exploit the model in deployment phase, we extract it from HDFS to the local file system.

B. Clustering model

We worked with a subset of dataset where observations have the "good" value for the fyAverage attribute. To avoid the problem of the dominance of large values of certain attributes, z-score standardization is applied to normalize continuous attributes.

We use stats package [23] which contains an implementation of k-means algorithm.

The k-means() function is invoked in mapreduce functions, and the model resulted has 5 clusters. Table VI shows the student's distribution according to department over the five clusters, where the recommended department has the highest percentage.

TABLE VI. STUDENT'S DISTRIBUTION OVER THE FIVE CLUSTERS

| cluster | GE | GI | GM | GP | TM |
|---------|--------|-------|-------|-------|-------|
| 1 | 13.45% | 8% | 67.3% | 4.55% | 6.64% |
| 2 | 57.08% | 12.5% | 3.96% | 7.71% | 18.7% |
| 3 | 4.85% | 64.6% | 4.3% | 3.08% | 23% |
| 4 | 19.62% | 21.5% | 15.8% | 35.2% | 7.76% |
| 5 | 2.28% | 23.2% | 15.1% | 10.4% | 48.8% |

VI. DEPLOYMENT

To deploy our models in the production environment, we develop a web application using JavaScript libraries as shown in Figure 2 and Figure 3. This application provides a dashboard with interactive visualization, and offers features which are divided in two main parts:

a) Exploratory data analysis

- To visualize the data stored in Hadoop cluster
- To summarize the main characteristics of students: proportion of sex, proportion of origins, etc.
- To visualize the evolution of student performance through years.

b) Predictive analysis

The decision-maker submits the dataset of the new enrollers, then he gets the result of the prediction for each department, and the recommended department.

ESTC Casablanca -RITM Laboratory

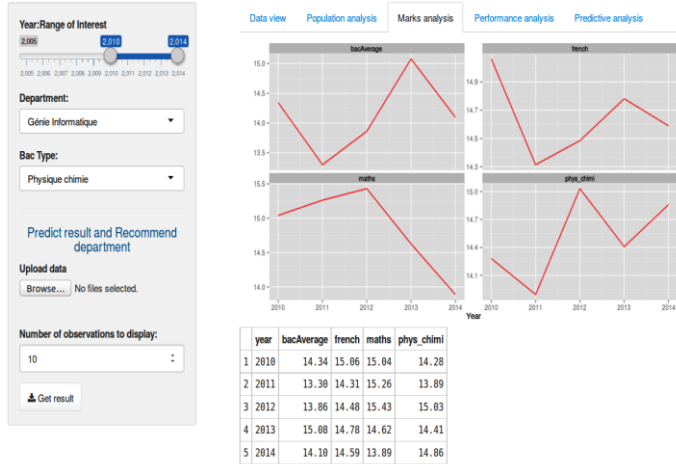


Figure 2. Marks analysis according to range of year, department and bac type

ESTC Casablanca -RITM Laboratory

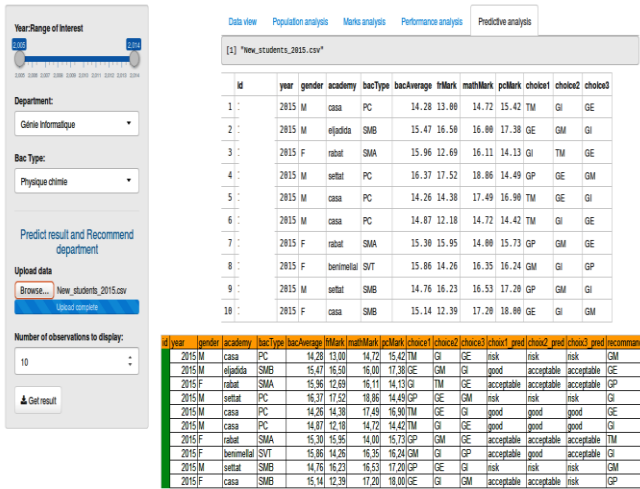


Figure 3. Predict result and recommend department

VII. CONCLUSION

The main goal of this work is to help to reduce the failure rate in the first year of a high school, and improve the enrollment management system. To realize that, we adopted R and Hadoop association to build a platform for analytics on massive data collected between 2005 and 2014. We built a classification model with C5.0 algorithm to predict the first year result for new students, and we also built a recommendation system by using k-means algorithm to recommend the suitable department. We finished by producing an interactive dashboard to visualize and deploy the finished models in the production environment. In the nearest future, we think that the use of this platform will

help us to study and analyze more complex data from various sources in the aim to analyze the real-time interaction between students and the environment where they learn.

REFERENCES

- [1] T. M. Mitchell. Machine Learning. McGraw-Hill, New York, 1997.
- [2] Quinlan, J.R. (1998) C5.0: An Informal Tutorial. <http://www.rulequest.com/see5-unix.html>.
- [3] Quinlan J R. , "C4.5: Programs for machine learning", Machine Learning, 1994(3): 235–240.
- [4] Corinna Cortes, Vladimir Vapnik, "Support-Vector Networks", Machine Learning, Vol. 20, pp. 273-297, 1995.
- [5] Brijesh Kumar Baradwaj and Saurabh Pal, "Mining Educational data to Analyze Students Performance", international journal of advanced computer science and applications, vol. 2, no. 6, 2011.
- [6] Zlatko J. Kovačić, "Early Prediction of Student Success: Mining Students Enrolment data ", research in higher education journal, vol. 15, p1, mar 2012.
- [7] M.Sindhuja et al., "Prediction and Analysis of Students Behavior using BARC algorithm",international journal on computer science and engineering (ijcse), vol. 5 no. 06 jun 2013.
- [8] Krina Parmar et al., "Prediction and Analysis of Student Performance using Distributed Data mining", international journal of emerging technologies and applications in engineering, technology and sciences (ij-eta-ets)december 2014 ,special issue.
- [9] Apache Hadoop.<https://hadoop.apache.org>
- [10] Sanjay Ghemawat et al., "The Google File System", ACM SIGOPS Operating Systems Review 2013.
- [11] Jeffrey Dean and Sanjay Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters", 2004.
- [12] <http://hadoop.apache.org/docs/r1.2.1/streaming.html#Hadoop+Streaming>
- [13] R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- [14] Revolution Analytics (2013). rmr2: R and Hadoop Streaming Connector. R package version 2.3.0.
- [15] Revolution Analytics (2013). rhdfs: R and Hadoop Distributed filesystem . R package version 1.0.8.
- [16] Hive . Ashish Thusoo et al., "Hive - a petabyte scale data warehouse using Hadoop", In Proceedings of International Conference on Data Engineering (ICDE), 2010, pages 996- 1005.
- [17] Hiveserver2:<https://cwiki.apache.org/confluence/display/Hive/HiveServer2+Clients>.
- [18] Apache Thrift. <http://thrift.apache.org/docs/concepts>.
- [19] NexR (2014). RHive: R and Hive. R package version 2.0-0.2. <http://CRAN.R-project.org/package=RHive>.
- [20] Max Kuhn, Steve Weston, Nathan Coulter and Mark Culp. C code for C5.0 by R. Quinlan (2015). C50: C5.0 Decision Trees and Rule-Based Models. R package version 0.1.0-24. <http://CRAN.R-project.org/package=C50>.
- [21] Freund Y, Schapire R E., "A decision-theoretic generalization of on-line learning and an application to boosting", Journal of Computer and System Sciences, 1997(1): 119–139.
- [22] Alexandros Karatzoglou, Alex Smola, Kurt Hornik, Achim Zeileis (2004). kernlab - An S4 Package for Kernel Methods in R. Journal of Statistical Software 11(9), 1-20. URL <http://www.jstatsoft.org/v11/i09/>.
- [23] stats. <https://stat.ethz.ch/R-manual/devel/library/stats/html/00Index.html>.

Leveraging DEMF to Ensure and Represent 5ws&1h in Digital Forensic Domain

Jasmin Cosic

ICT Section of Police Administration
Ministry of the Interior of Una-Sana canton
Bihac, Bosnia and Herzegovina

Miroslav Baca

Faculty of Organization and Informatics
University of Zagreb
Varazdin, Croatia

Abstract— in this paper authors will discuss about metadata of digital evidence, digital chain of custody and framework for digital evidence management. A concretization of DEMF - Digital Evidence Management Framework will be presented. DEMF is a framework that was developed on conceptual – ontology level, and now programmed in java.

One case study from real criminal case through DEMF will be presented.

Keywords – digital evidence management framework, DEMF, intererability, digital evidence, chain of custody

I. INTRODUCTION

The concept “chain of custody” is not new in digital forensic domain. There is a lot of published scientific paper and research about this concept in digital forensic domain too. This concept was developed in medicine, engineering, construction, criminal law, and in each domain where it is necessary to have a chronological consistency of some events with accompanying metadata. A similar situation exists in the field of digital forensics. Within the domain of digital forensics, in the digital evidence management, chain of custody includes documentation of everything, the chronological consistency of “events” related to digital evidence. If that chain of custody concept is digital, in digital world, and if in addition to all metadata, 5ws&1h can assure the integrity of digital evidence - that will be a big step forward.

In this paper, the authors will analyse this issue, describe DEMF¹ using ontology and OWL, and make one case study from a real life. In this case study, authors will simulate an

actual case, in which the "first responders" in particular digital investigation found three digital evidence. They were added to the ".demf" container using DEMF applications. After that, complete container, containing the original digital evidence and metadata, was encrypted with AES256 method, and was delivered to another participant in this chain - the investigator. He re-analysed the entire case again, "processed" these three digital evidence, added another file – new digital evidence into the container, and packed again everything in a new container – “.demf” and secured the container with a AES256 key.

Court expert, hired by the court for this case, received the court order, as well as the container with the necessary key for decryption. He checked the files, which, in this case are digital evidence, compared the hash values, as well as time stamps, geo-information, and other metadata that he secured through the DEMF. After completion of the findings, report and opinion, the complete case was returned to the court for further proceedings.

II. DIGITAL CHAIN OF CUSTODY

Term “Chain of preservation” or “chain of custody” according to [1] refers to a complete audit and control of original evidence material that could potentially be used in legal purposes. According to NIJ² [2] chain of custody is a process that maintains and documents chronological history of evidence (Document must include all data, name and surname, and/or initials of person who collected evidence, every person or entity who had access to evidence, date of collecting the evidence or their location change, name of the agency and case number, name of the victim or suspect - detailed description of all).

¹ DEMF stands for Digital Evidence Management Framework (Cosic&Baca, 2010)

² NIJ - National Institute of Justice (US National Institute of Justice)

Today, some authors use term "chain of evidence" rather than "chain of custody". The purpose of testimony about the chain of evidence is to prove that digital evidence has not been changed at any stage of forensic investigation and must include documentation on how it was collected, where it was collected, how was it transported, analysed and stored. Interruption of chain of custody leads to suspicion that evidence was changed, falsified or replaced. It is not enough for court to know only the exact location of digital evidence, but entire way of movement for the whole time must be recorded. Record and strict control of access to digital evidence also must be kept [3]–[11].

III. ONTOLOGY DESCRIPTION OF DEMF

For these reasons, the authors [12] [10] developed DEMF-Digital Evidence Management Framework, a conceptual framework that allowed not only for digital evidence to be chronologically followed in the chain of custody in order to monitor what, who, when, how, why and where came into contact with digital evidence, but it also secured preservation of the integrity of digital evidence through the use of hash functions and AES256 encryption.

In their early work [13] authors have tried to present DEMF by using ontology and OWL in order to present the knowledge in the domain of digital forensics and digital evidence, and "reusability" as well. Ontology was subsequently expanded, and apart from class, properties, and axiom, instances were created as well as rules that enable reasoning and help for individuals who make decisions on acceptability (judges, etc...).

Created ontology is published online at the OntoHub at following link: http://ontohub.org/demf/DEMF_V1.owl. During ontology creation process, authors are especially careful on reusability, and therefore they include in DEMF a following ontology: "Small Scale Digital Devices Ontology", "Digital Evidence Ontology" and "Cyber Forensic Ontology" [14][15][13]

IV. USING DEMF IN REAL ENVIRONMENT

Figure 1 presents a class diagram of typical activities in digital investigation process. These activities describe a digital investigation process, and would be essential for our example (study).

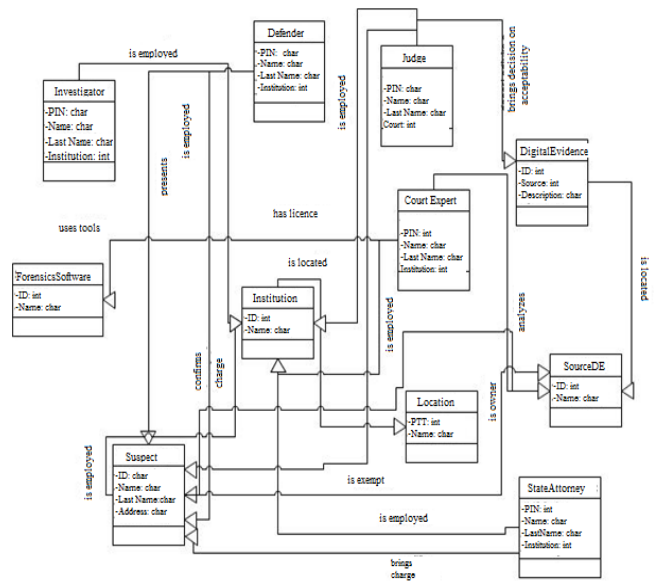


Figure 1 UML Class diagram with joining of main activities in DI process

Created DEMF ontology enables to create specific instances with the properties, take advantage of "reasoner" and tuned finished rules in the ontology in order to accurately determine which digital evidence is acceptable and which is not. Any digital evidence that meets the requirements of 5ws&1h will be "formally" acceptable.

For the purpose of case study, we will enter some data from a particular case into the DEMF application and then check the application outputs:

```
[Tue Nov 10 10:17:19 CET 2015] User logged in DEMF with ID = JASMIN COSIC from IP address = 31.176.223.
[Tue Nov 10 10:17:19 CET 2015] Working directory: C:\JRE_DEMF (x86)
[Tue Nov 10 10:17:25 CET 2015] Case CrimeSceneBihac11102015 loaded in memory.
[Tue Nov 10 10:17:43 CET 2015] Getting a hash of file ..
[Tue Nov 10 10:17:43 CET 2015] Hash file is retrieved.
[Tue Nov 10 10:17:43 CET 2015] Reading a meta-data ..
[Tue Nov 10 10:17:44 CET 2015] Meta-data of file is retrieved = Name=B.Form_1b_-_PhD2015.doc cp:revision=4 date=2015-06-02T10:52:00Z Company= Keywords= meta:word-count=643 subject= dc:creator=pc7 extended-properties:Company= meta:print-date=2015-06-02T10:49:00Z Word-Count=643 dcterms:created=2015-06-02T10:49:00Z dcterms:modified=2015-06-02T10:52:00Z Last-Modified=2015-06-02T10:52:00Z title= Last-Save-Date=2015-06-02T10:52:00Z meta:character-count=3668 Template=Normal.dotm meta:save-date=2015-06-02T10:52:00Z dc:title= Application-Name=Microsoft Office Word modified=2015-06-02T10:52:00Z Edit-Time=1800000000 cp:subject= Content-Type=application/msword X-Parsed-By=org.apache.tika.parser.DefaultParser X-Parsed-By=org.apache.tika.parser.microsoft.OfficeParser creator=pc7
```

`meta:author=pc7 dc:subject= extended-
properties:Application=Microsoft Office Word meta:creation-
date=2015-06-02T10:49:00Z Last-Printed=2015-06-02T10:49:00Z
meta:last-author= Comments= Creation-Date=2015-
06-02T10:49:00Z xmpTPg:NPages=2 Last-Author=
w:comments= Character Count=3668 Page-Count=2 Revision-
Number=4 extended-properties:Template=Normal.dotm
meta:keyword= Author=pc7 comment= meta:page-count=2`

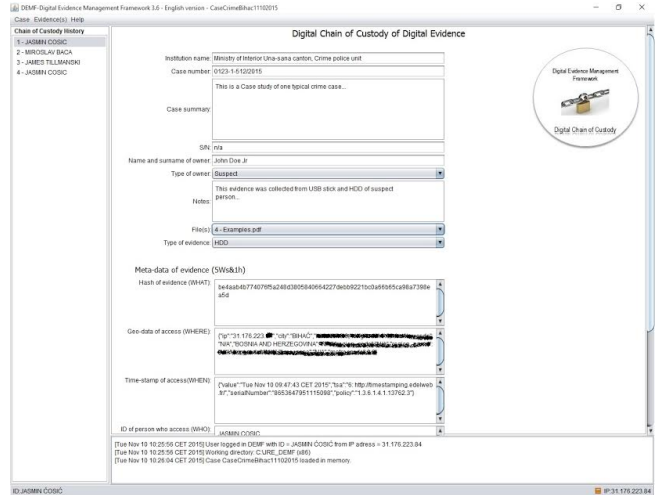


Figure 2 The complete history of accessing to digital evidence

In the aforementioned log file, we can find generated metadata, which describe digital evidence and application: User ID, IP address, Digital evidence hash value, Place and geo-data of evidence, and finally a Time stamp. Court order – the Reason of access and Procedures are fixed variables that we entering during the first use of the application.

Timestamp that we get in DEMF are from trusted time stamping services (obtained from internet), and geo-data are also obtained from web services. This can be configurable and we can have up to 5 different servers.

On next cycle, in the further use of the “.demf” container, if the integrity of digital evidence were compromised, DEMF would give an alarm alert - as following:

`[Tue Nov 10 10:23:21 CET 2015] WRONG HASH.
Existing =
3c55d29dfcda36f97293f4104c64d36a252242bbd554a953d762
2be7fce93e4c ,
Current =
f621bb81677242fc19159b9d892e734b593c9ca9db7be8d9777
91939c2746e55`

Considering that digital evidence, like any other digital file, have a life cycle [16], every new access to container that generates DEMF (.demf files), shall preserve a complete history of 5WS & 1H. At the end of the investigation process, when the evidence must be presented to the court, investigators will have uninterrupted and preserved chain of custody. In Figure 2, we can see the whole access history to digital evidence in a particular case.

Now, if we define acceptability of digital evidence as[17]:

```
OriginalDigitalEvidence (?x) ^ hasFingerprint (?x,  
?fp) ^ hasHashValue (?x, ?sha2) ^  
hasEvidenceLocationCoordinates (?x, ?gps) ^  
hasEvidenceProcedure (?x, ?proc) ^  
hasAccessReason (?x, ?reas) ^ hasAccessTime (?x, ?ts)  
→ DigitalEvidenceAcceptable (?x)
```

We can say that such digital evidence is formally acceptable.

Why formally acceptable? Because it is ultimately the last decision of the court (judge or jury) whether the evidence will be admissible or not admissible!

Judges decide what evidence will or will not be allowed in their courtrooms. How do judges make this decision - according to Fry, Daubert principle, stochastic or in another way, is especially question!

Today, there is a very few authors deals with the problem of acceptability of digital evidence [17] [18]

CONCLUSION AND FURTHER RESEARCH

The problem related to DEMF is a problem that is typical for any other case management systems - interoperability and data exchange with other systems.

There is no adopted standard and procedures for the exchange and interoperability of information/data between different agencies and countries. Every agency that wants to "read" for example, data from “.demf” containers must have a secret key, and of course installed DEMF application or at least DEMF reader.

Author [19] in his work summarizes the strengths and weaknesses of existing digital evidence schemas, and proposes the open-source CybOX schema as a foundation for storing and sharing digital forensic information. An additional open-source schema and associated ontology called Digital Forensic Analysis eXpression (DFAX) is proposed that provides a layer of domain specific information overlaid on CybOX.

In future research, the authors will deal with this problem, present DEMF through Cybox and DFAX and try to fill this gap by representing the metadata in the file “.demf” using the proposed DFAX standard.

ACKNOWLEDGMENT

Research and development of DEMF in the first phase was supported by the Centre for biometrics, forensics and privacy, Faculty of Organization and Informatics, University of Zagreb. I would also like to thank my friend, software engineer, Edo Civic, who assisted me to develop some java parts of DEMF and DEMF presenter.

REFERENCES

- [1] R. Yaeger, “Criminal computer forensics management,” in *Proceedings of the 3rd annual conference on Information security curriculum development InfoSecCD 06*, 2006, p. 168.
- [2] National Institute of Justice, “Crime Scene Investigation: Guides for Law Enforcement,” 2011. [Online]. Available: <http://www.nij.gov/topics/law-enforcement/investigations/crime-scene/guides/glossary.htm>.
- [3] M. Nagaraja, “Investigators Chain of Custody in Digital Evidence Recovery,” *Indian Police Serv.*, pp. 1–7.
- [4] SWGIT, “Best Practices for Maintaining the Integrity of Digital Images and Digital Video,” 2007.
- [5] National center for Forensic, “Digital evidence in the courtroom: A guide for preparing digital evidence for court presentation.” 2003.
- [6] B. Pladna, “Computer Forensics Procedures, Tools, and Digital Evidence Bags: What They Are and Who Should Use Them,” in *Computer Forensics Procedures, Tools, and Digital Evidence Bags 3*, 2009, pp. 1–15.
- [7] J. Patzakis, “Maintaining The Digital Chain of Custody,” in *IFOSEC*, 2003.
- [8] G. Giova, “Improving Chain of Custody in Forensic Investigation of Electronic Digital Systems,” *Int. J. Comput. Sci. Netw. Secur.*, vol. 11, no. 1, 2011.
- [9] E. Casey, *Digital evidence and computer crime: forensic science, computers and the Internet*. Academic Press, 2004.
- [10] J. Cosic and M. Bača, “Do We Have Full Control Over Integrity in Digital Evidence Life Cycle?,” in *32nd International Conference on Information Technology Interfaces (ITI)*, 2010, pp. 429–434.

- [11] C. L. T. Brown, *Computer evidence: Collection and Preservation*. Course Technology PTR, 2009.
- [12] J. Cosic and M. Bača, “A Framework to (Im)Prove „Chain of Custody“ in Digital Investigation Process,” in *Proceedings of the 21st Central European Conference on Information and Intelligent Systems*, 2010, no. Im.
- [13] J. Čosić, Z. Čosić, and M. Bača, “An ontological approach to study and manage digital chain of custody of digital evidence,” *J. Inf. Organ. ...*, 2011.
- [14] A. Brinson, A. Robinson, and M. Rogers, “A cyber forensics ontology: Creating a new approach to studying cyber forensics,” *Digit. Investig.*, vol. 3, pp. 37–43, Sep. 2006.
- [15] D. C. Harrill and R. P. Mislán, “A Small Scale Digital Device Forensics ontology,” *Small Scale Digit. Device Forensics J.*, vol. 1, no. 1, pp. 1–7, 2007.
- [16] J. Cosic and Z. Cosic, “Chain of Custody and Life Cycle of Digital Evidence,” *Comput. Technol. Appl.*, vol. 3, pp. 126–129, 2012.
- [17] J. Čosić, “Izgradnja otvorenog okvira za uspostavu i očuvanje lanca dokaza u forenzičkoj analizi digitalnih dokaza,” PhD thesis, University of Zagreb, 2014.
- [18] G. C. Kessler, “Judges’ Awareness, Understanding, and Application of Digital Evidence by,” Graduate School of Computer and Information Science Nova Southeastern University, 2010.
- [19] E. Casey, G. Back, and S. Barnum, “Leveraging CybOX™ to standardize representation and exchange of digital forensic information,” *Digit. Investig.*, vol. 12[1] E. C, pp. S102–S110, 2015.

AUTHORS PROFILE

Doc.dr.sc. Jasmin Čosić has received his Doctoral degree in Information Science, in digital forensic domain, from University of Zagreb (Croatia) in 2014. He is working in Ministry of the Interior of Una-sana canton in B&H in ICT Section. He is a Assistant professor (docent) in Pedagogical Faculty (Department for mathematics and informatics) and Biotechnical Faculty – University in Bihac (B&H). He is also a ICT Expert Witness, and a member of various professional societies. His areas of interests are Digital Forensic, Computer Crime, Biometrics, Information Security and DBM Systems.
He is author or co-author more than 40 scientific papers and one book.

Prof.dr.sc. Miroslav Bača is currently an Full professor, University of Zagreb, Faculty of Organization and Informatics. He is a member of various professional societies and program committee members, and he is reviewer of several international journals and conferences. He is also the Head of the Biometrics centre in Varaždin, Croatia. He is author or co-author more than 100 scientific and professional papers and two books. His main research fields are Computer Forensics, Biometrics and Privacy.

GCTrie for Efficient Querying in Cheminformatics

Yu Wai Hlaing

Ph.D candidate

University of Computer Studies
Yangon, Myanmar

Kyaw May Oo

Faculty of Computing

University of Information Technology
Yangon, Myanmar

Abstract— *The field of graph indexing and query processing has received a lot of attention due to the constantly increasing usage of graph data structures for representing data in different application domains. To support efficient querying technique is a key issue in all graph based application. In this paper, we propose an index trie structure (GCTrie) that is constructed with our proposed graph representative structure called graph code. Our proposed GCTrie can support all types of graph query. In this paper, we focus on index construction, subgraph query and supergraph query processing. The experimental results and comparisons offer a positive response to the proposed approach.*

Keywords—*graph indexing and querying; graph representative structure; index; subgraph query; supergraph query*

I. INTRODUCTION

Many scientific and commercial applications urge for patterns that are more complex and complicated to process than frequent item sets and sequential patterns. Such sophisticated patterns range from sets and sequences to trees, lattices and graphs. As one of the most general form of data representation, graphs easily represent entities, their attributes and their relationships to other entities. The significant of using graphs to represent complex datasets has been recognized in different disciplines such as chemical domain [6], computer vision [7], and image and object retrieval [8]. Various conferences over the past few years on mining graphs have motivated researchers to focus on the importance of mining graph data. Different applications result in different kinds of graphs, and the corresponding challenges are also quite different. A graph describes relationships over a set of entities. With nodes and edges labels, a graph can depict the attributes of both the entity set and the relation. For example, chemical data graphs are relatively small but the labels on different nodes (which are drawn from a limited set of elements) may be repeated many times in a single molecule (graph).

Storing the graphs into large datasets is a challenging task as it deals with efficient space and time management. Over the years, a number of different representative structures have been developed to represent graphs more and more efficiently and uniquely. Developing such structures is particularly challenging in terms of storage space and generation time. Among many representative structures adjacency list [9] and adjacency matrix [10] are the most common. We have already proposed a new graph representative structure called graph code [1]. Graph code is a new way of representing graphs to support all kinds of graph queries without verifying between

graph structures. A good graph indexing and querying approach should have compact indexing structures and has a good power of pruning the false graphs in the dataset. The strategy of graph indexing is to move high costly online query processing to off-line index construction phase [2]. Chemical graphs in datasets are undirected labelled graphs. So, graph code is developed to process undirected labelled graphs. Graph code can retain the structural information of original graph such as which two edges are connected on which vertex.

To effectively understand and utilize any collection of graphs, an approach that efficiently supports elementary querying mechanism is crucially required. Given a query graph, the task of retrieving related graphs as a result of the query from a large graph dataset is a key issue in all graph based applications. This has raised a crucial need for efficient graph indexing and querying approaches. A primary challenge in computing the answers of graph queries is that pair-wise comparisons of graphs are usually really hard problems. It is apparent that the success of any graph based application is directly dependent on the efficiency of the graph indexing and query processing mechanisms. Recently, there are many techniques that have been proposed to tackle these problems.

In principle, queries in graph datasets can be broadly classified into the four categories: graph isomorphism query, subgraph query, supergraph query, and similarity query. Most of the existing graph indexing and querying approaches proposed to deal with only one type of the query problem. Our proposed approach allows the chemical compound dataset to be queried chemical structures in terms of XML file format. Using proposed approach, all types of graph queries can be processed. After entering a chemical structure as a query, user can process their desired query types. In this paper, we describe our proposed graph code structure, and GCTrie, and also perform subgraph query and supergraph query processing by probing GCTrie. We also perform experimental analysis on index construction and on these queries using proposed approach and other existing approaches.

II. PRELIMINARIES

For simplicity, we present the key concepts, notations, and terminology used in our proposed approach which includes labeled undirected graph, graph automorphism, subgraph query, supergraph query, and graph code.

As a general data structure, labeled graphs is used to model complicated structures and schemaless data. In labeled graph, vertex and edge represent entity and relationship, respectively.

The attributes associated with entities and relationships are called labels. XML is a kind of directed labeled graph. The chemical compound shown in Fig. 1 is labeled undirected graph.

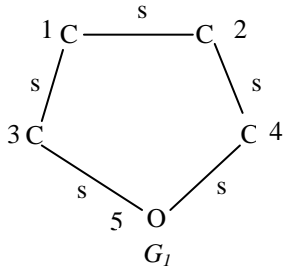


Figure 1. A Labeled Undirected Graph

Definition 1. Labeled Undirected Graph

A labeled undirected graph G is defined as 5-tuple, (V, E, L_V, L_E, l) where V is the non-empty finite vertex set called vertices, and E is the unordered pairs of vertices called edges. L_V and L_E are the set of labels of vertices and edges and l is a labelling function assigning a label to a vertex $l: V$ to L_V and an edge $l: E$ to L_E .

Definition 2. Graph Isomorphism

Let $G = (V, E, L_V, L_E, l)$ and $G' = (V', E', L'_V, L'_E, l')$ be two graphs. An automorphism between two graphs G and G' is an isomorphism mapping where $G = G'$. An isomorphism mapping is a mapping of the vertices of G to vertices of G' that preserve the edge structure of the graphs. That is, it is a graph isomorphism from a graph G to itself.

Definition 3. Subgraph Query

This category searches for a specific pattern in the graph dataset. The pattern can be a small graph. Therefore, given a graph dataset $D = \{G_1, G_2, \dots, G_i\}$ and a subgraph query q , the answer set $A = \{G_i | q \subseteq G_i, G_i \in D\}$.

Definition 4. Supergraph Query

Given a graph dataset $D = \{G_1, G_2, \dots, G_i\}$ and a supergraph query q , if a query q is a supergraph of a dataset graph, the answer set $A = \{G_i | G_i \subseteq q, G_i \in D\}$.

Definition 5. Graph Code

For a graph G_i , the code of G_i , denoted by $c(G_i)$ is the list of the form $e_{id}[v, e_{id_adj}] \dots$ depending on adjacent edges. e_{id} is the edge id, v is vertex label on which two edges are connected, e_{id_adj} is adjacent edge id for this edge.

III. RELATED WORKS

Graphs are used to represent many real life applications. Graphs can be used to represent networks. The networks may include paths in a city or telephone network or circuit network. Graphs are also used in social networks like linkedIn, facebook. Many graph datasets (e.g., chemical compounds) have more than one vertex with the same label. Same graph is stored more than once in the graph datasets leading to adverse results of mining. To ensure the consistency of graph datasets,

required a mechanism to check whether two graphs are automorphic or not. So, detection and elimination of automorphic graphs is required. In proposed approach, a graph is represented via its graph code generated by using adjacent edge information and edge dictionary. Instead of expensive graph automorphism test, automorphic graphs can be detected by matching codes of two graphs [3].

GraphGrep [4] was proposed that is a path-based technique to index graph datasets. It has three basic components: building the index to represent graphs as sets of paths, filtering dataset based on query and computing exact matching. GraphGrep enumerates paths up to a threshold length (l_p) from each graph. An index table is constructed and each entry in the table is the number of occurrences of the path in the graph. Filtering phase generates a set of candidate graphs for which the count of each path is at least that of the query. Verification phase verifies each candidate graph by subgraph matching. However, the graph dataset contains huge amount of paths and can have an effect on the performance of the index.

OrientDB [5] is an open source NoSQL database management system written in java. It is a multi-model database, supporting graph, document, key/value, and object models, but the relationships are managed as in graph databases with direct connections between records. It supports schema-less, schema-full and schema-mixed modes. It has a strong security profiling system based on users and roles and supports querying with SQL extended for graph traversal.

IV. PROPOSED APPROACH

In our proposed approach, there are three main phases: code generation phase, subgraph query and supergraph query processing phase, and graph isomorphism query and similarity query processing phase. There are three sub-steps in code generation phase. These are preprocessing, code generation and automorphism checking, and index construction. In subgraph query and supergraph query processing phase, there are four sub-steps: preprocessing, code generation, subgraph querying and supergraph querying. In graph isomorphism query and similarity query processing phase, there are also four sub-steps, preprocessing, code generation, and graph isomorphism querying and similarity querying. In this paper, we focus on index construction step, and subgraph query and supergraph query processing phase.

A. Preprocessing, CodeGeneration and Automorphism Checking

In preprocessing, the graph information such as vertex information, edge information, and adjacent edge information are generated by parsing input xml files with xml parser. The edge information of the graph is defined as (V_{id}, L, V_{id}) where V_{id} is the vertex id, L is the edge label. Then adjacent edge information is generated. Fig. 2 shows graph information for graph G_1 in Fig. 1.

| | | | | | |
|-------------------------------------|------------------------|---------|---------|---------|---------|
| Vertex id : | 1 | 2 | 3 | 4 | 5 |
| Vertex Info : | C | C | C | C | O |
| (a) | | | | | |
| V_{ib}, L, V_i: | <1,s,2> | <1,s,3> | <2,s,4> | <3,s,5> | <4,s,5> |
| Edge Info: | <C,s,C> | <C,s,C> | <C,s,C> | <C,s,O> | <C,s,O> |
| (b) | | | | | |
| Edge: | Adjacent Edges: | | | | |
| <1,s,2> | <1,s,3>, <2,s,4> | | | | |
| <1,s,3> | <1,s,2>, <3,s,5> | | | | |
| <2,s,4> | <1,s,2>, <4,s,5> | | | | |
| <3,s,5> | <1,s,3>, <4,s,5> | | | | |
| <4,s,5> | <2,s,4>, <3,s,5> | | | | |
| (c) | | | | | |

Figure 2. Graph Information of G_i (a) Vertex Information (b) Edge Information (c) Adjacent Edge Information

For each edge from the graph's edge information, check the edge dictionary to determine whether the edge is already existed in edge dictionary or not. If not, insert new edge into edge dictionary. Then, the edge ids are associated with their corresponding edges in graph's edge information. Edge dictionary is shown in Fig. 3.

| Id | Edge |
|-----|---------|
| 1 | <C,s,C> |
| 2 | <C,s,O> |
| ... | ... |

Figure 3. Edge Dictionary

A graph is represented holistically into a graph code that preserves the structural information of the graph. Every edge in the graph is assigned with global unique identifier already defined in the edge dictionary. Instead of using the edge itself, using the edge id of the edge dictionary can have advantages in three ways:

- Firstly, using the edge id in the code saves the amount of storage space.
- Secondly, using the same id for the duplicated edge is effective when constructing the graph code.
- Thirdly, using the edge id in the code reduces the time for finding automorphic or isomorphic graphs.

Most of the chemical graphs have a lot of common edges. So, edge dictionary uses little memory space. Edge dictionary and adjacent edge information are used to generate graph code. Graph code for graph G_i is as follows:

$$c(G_i) = 1[c,1], 1[c,1], 1[c,1], 1[c,2], 1[c,1], 1[c,2], 2[c,1], 2[o,2], 2[c,1], 2[o,2]$$

After computing the graph code of G_k , compares it with each graph code G_i in code store (CS), $1 \leq i < k$, to check graph automorphism. If the graph code of G_k has the same code as that of G_i , concludes that the two graphs are automorphic and append id of G_k to corresponding graph code of G_i . Otherwise, add the graph code of G_k to CS assuming as G_k is a new graph.

B. Index Construction

After generating graph codes for all dataset graphs and checking automorphism, the next step is to construct GCTrie for efficient querying. Instead of using path or subgraph decomposition to support subgraph query type which has result in structural information lost and exhaustive enumeration time problems, we propose an index trie structure called GCTrie for supporting all types of graph query. We put the graph codes of all dataset graphs in GCTrie. A GCTrie is a trie where each node except the root node is a string array that represents an edge id or a vertex label on which two edges are connected. There are five levels in the GCTrie. The second and fourth level is for edge ids. The third level is for vertex labels and the last is for leaves which are implemented by hashmaps of graph ids and their frequencies. Procedure for index construction is shown as follows. We represent one edge's adjacent code, e.g; $1[o,2]$ as feature f .

Procedure. IndexConstruction(CS)

```

For each  $c(G_i) \in CS$ 
  For each feature  $f \in c(G_i)$ 
    Put  $f$  in GCTrie
return GCTrie

```

In index GCTrie construction, for the graph code of G_i from CS, $c(G_i)$: $1[c,1], 1[c,1], 1[c,1], 1[c,2], 1[c,1], 1[c,2], 2[c,1], 2[o,2], 2[c,1], 2[o,2]$, there are four occurrences of features $1[c,1]$. There are two occurrences of features $1[c,2]$, $2[c,1]$ and $2[o,2]$. So, the GCTrie after putting $c(G_i)$ is as shown in Fig. 4.

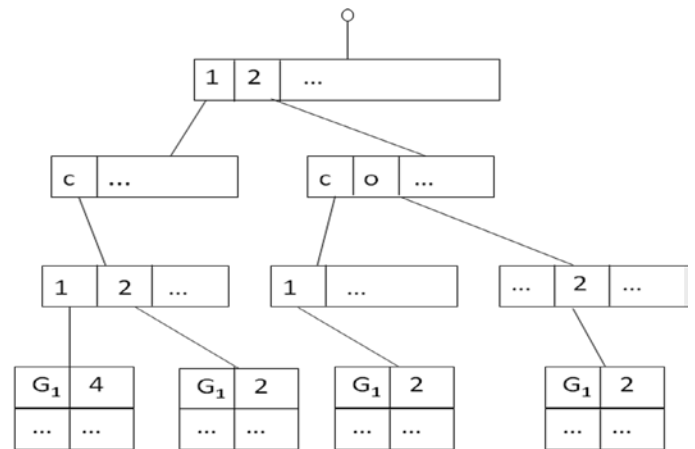


Figure 4. GCTrie After Putting $c(G_i)$

Then we put all dataset graph codes from CS into this GCTrie. When the query graph enters into the system, vertex information, edge information and adjacent edge information of query graph is generated in preprocessing step. Then the query graph code is generated in code generation phase using system's edge dictionary and query adjacent edge information. Each feature from query graph code is probed in GCTrie.

C. Subgraph Querying

In the core of many graph-related applications lies a common and critical problem of how to efficiently process subgraph query. In some cases, the success of an application directly relies on the efficiency of the query processing system. The classical graph query problem can be described as follows: given a graph dataset $D = \{G_1, G_2, \dots, G_n\}$ and a graph query q , finds all the graphs in which q is a subgraph. If all of the query features are matched with features of data graph codes in the GCTrie but the size of dataset graphs are larger than query graph size, then the dataset graphs are returned as answer set. The following algorithm 1 describes the step-by-step process for subgraph query.

Algorithm 1 SubgraphQuery

Input : GCTrie, and Query q

Output : Answer set D_q .

1. Generate graph code for query q .
 2. Let $D_q = D$
 3. For each feature $qf \in c(q)$
 4. Probe qf in GCTrie.
 5. If $qf \in \text{GCTrie}$
 6. Intersect D_q and D_{qf} .
 7. For each $G_i \in D_q$
 8. If $\text{size}(G_i) < \text{size}(q)$
 9. Remove G_i from D_q .
 10. Return D_q ;
-

Assume that we have generated the graph code of the query graph. We establish a necessary condition that forms the basis for processing subgraph query. Thus we state the following theorem.

Theorem 1 Given a query graph q , if q is a subgraph of a dataset graph G , then $c(q) \subseteq c(G)$.

Proof. By definition, if q is a subgraph of G , then every feature of q appears in G . Therefore, if parametric quantities of $c(q)$ are contained in $c(G)$, then $c(q) \subseteq c(G)$.

The intuition is as follows. If a query q is a subgraph of a dataset graph, then all of its features are a subset of the features of the dataset graph. Therefore, the adjacent edges of each edge that appear in the graph code of the query will definitely appear in the graph code of the dataset graph.

D. Supergraph Querying

Supergraph query searches for the graph dataset members of which their whole features are contained in the input query. Formally, given a dataset $D = \{G_1, G_2, \dots, G_n\}$ and a supergraph query q , if q is a supergraph of the dataset graphs then all of its features form a superset of the features of the resulted dataset

graphs. The large number of graphs in datasets and the NP-completeness of subgraph isomorphism testing make it challenging to efficiently processing supergraph queries.

In our propose approach, when the query graph enters, it is represented as a query graph code. Each feature from query' graph code is probed in GCTrie. If all of the query features are matched with features data graph codes in the GCTrie but the query graph size is larger than the dataset graphs' size. Then the dataset graphs are returned as answer set that are contained in query as subgraph. The step-by-step process of supergraph query is described as the following algorithm 2.

Algorithm 2 SupergraphQuery

Input : GCTrie, and Query q

Output : Answer set D_q .

1. Generate graph code for query q .
 2. Let $D_q = D$
 3. For each feature $qf \in c(q)$
 4. Probe qf in GCTrie.
 5. If $qf \in \text{GCTrie}$
 6. Intersect D_q and D_{qf} .
 7. For each $G_i \in D_q$
 8. If $\text{size}(G_i) > \text{size}(q)$
 9. Remove G_i from D_q .
 10. Return D_q ;
-

Assume that we have generated the graph code of the query graph. We establish a necessary condition that forms the basis for processing supergraph query. Then we state the following theorem.

Theorem 2 Given a query graph q , if q is a supergraph of a dataset graph G , then $c(G) \subseteq c(q)$.

Proof. By definition, if q is a supergraph of G , then every feature of G appears in q . Therefore, if parametric quantities of $c(G)$ are contained in $c(q)$, then $c(G) \subseteq c(q)$.

The intuition is as follows. If a query q is a supergraph of a dataset graph, then all of its features form a superset of the feature of the resulted dataset graphs. Therefore, the adjacent edges of each edge that appear in the graph code of the dataset graph will definitely appear in the graph code of the query.

V. EXPERIMENTAL ANALYSIS

A performance analysis for proposed approach is presented in this section. The main goal of the experiment is to represent the performance evaluation of our proposed approach apply on AIDS antiviral screen compound dataset, NCI yeast anticancer drug screen dataset, and primary screening dataset for Formylpeptide Receptor.

Index Construction times for graph indexing approaches such as GraphGrep, OrientDB and our proposed approach are analyzed. All of the approaches are implemented in java on Intel(R) Core (TM) i3-4010U CPU with 2GB memory and Window7 34-bit operating system. Fig. 5, Fig. 6, and Fig. 7

shows the index construction times vary for different approaches on three datasets respectively. Various numbers of chemical graphs are tested and take the average index construction times to compare GraphGrep, OrientDB and our proposed approach. For GraphGrep, We use two values 4 and 10 for parameter: the length of path (l_p). It can be seen that our proposed approach consumes at least 10 times less than OrientDB in index construction and at least 10^2 times less than GraphGrep ($l_p = 4$) and GraphGrep ($l_p = 10$) respectively.

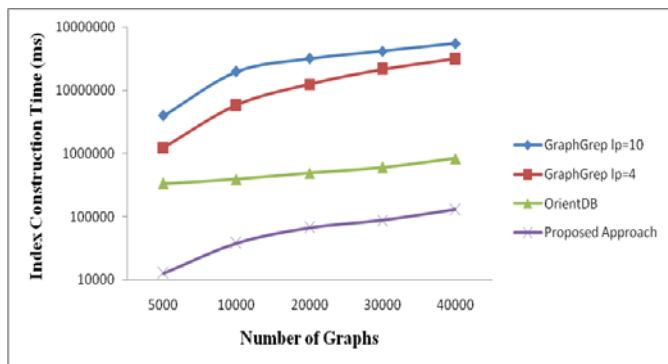


Figure 5. Analysis of Index Construction Time of Three Different Approaches for AIDS Antiviral Screen Dataset

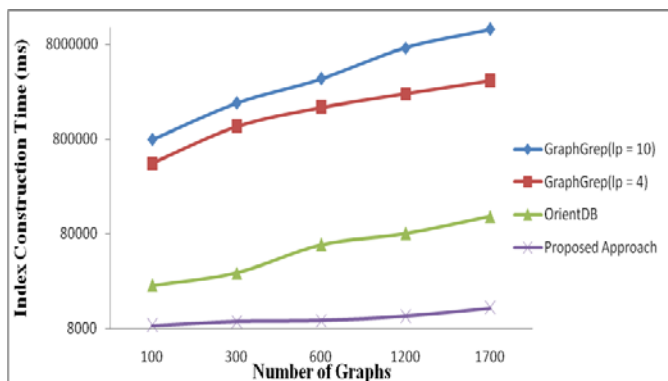


Figure 6. Analysis of Index Construction Time of Three Different Approaches for NCI Yeast Anti-cancer Drug Screen Dataset

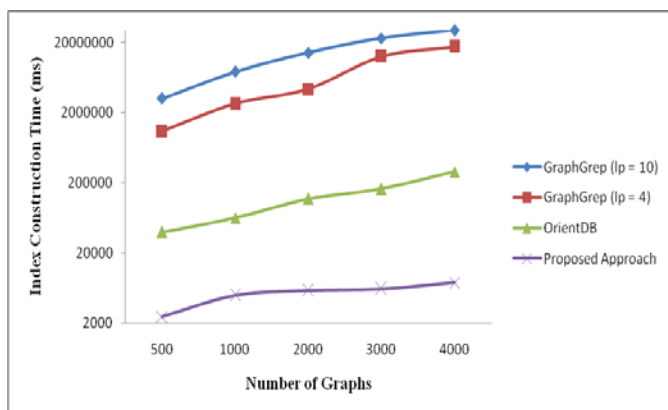


Figure 7. Analysis of Index Construction Time of Three Different Approaches for Primary Screening Dataset for Formylpeptide Receptor

We evaluate the query response time of our proposed approach with GraphGrep on AIDS antiviral screen dataset, NCI yeast anti-cancer drug screen dataset and primary screening dataset for Formylpeptide Receptor. Since GraphGrep only support subgraph isomorphism query, we can evaluate subgraph isomorphism query response time with it. For GraphGrep, we use two values 4 and 10 for parameter; the length of path (l_p). Fig. 8 shows the analysis of subgraph isomorphism query response time over AIDS antiviral screen dataset. It can be seen that our proposed approach significantly reduces at least 10^3 times for subgraph isomorphism query response time when compare to GraphGrep.

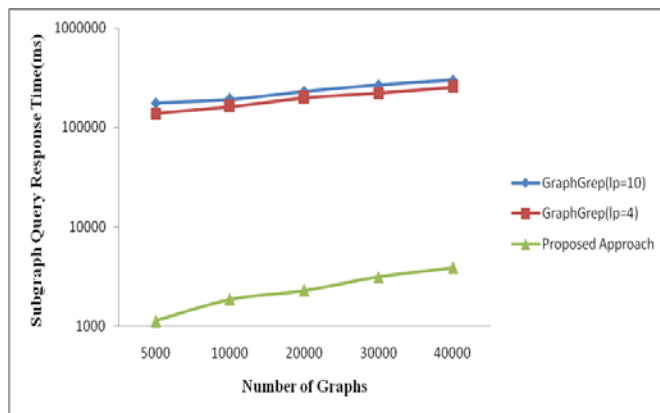


Figure 8. Analysis of Subgraph Query Response Time Between GraphGrep and Proposed Approach on AIDS antiviral screen Dataset

Fig. 9 shows the analysis of subgraph isomorphism query response time over NCI yeast anti-cancer drug screen dataset. Fig. 10 shows the analysis of subgraph isomorphism query response time over primary screening dataset for Formylpeptide Receptor. It can be seen that our proposed approach significantly reduces at least 10^2 times and 10^3 times of subgraph isomorphism query response time less than when compare to GraphGrep over NCI yeast anti-cancer drug screen dataset and primary screening dataset for Formylpeptide Receptor respectively.

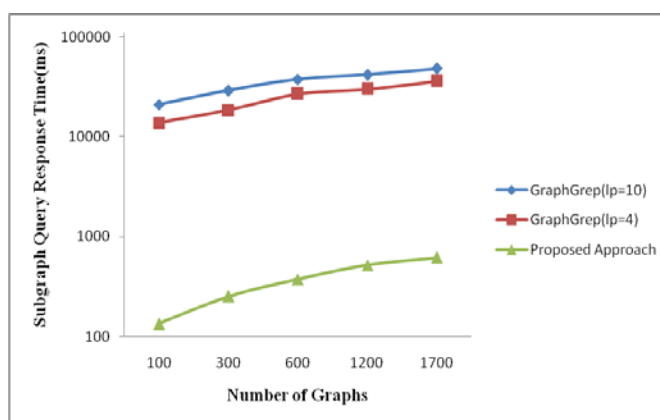


Figure 9. Analysis of Subgraph Query Response Time Between GraphGrep and Proposed Approach on NCI Yeast Anti-cancer Drug Screen Dataset

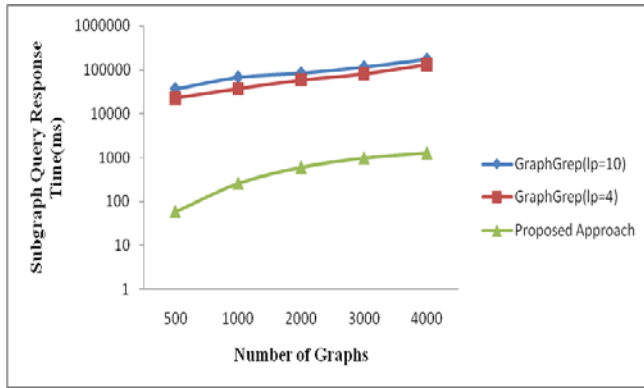


Figure 10. Analysis of Subgraph Query Response Time Between GraphGrep and Proposed Approach on Primary Screenin Dataset for Formylpeptide Receptor Dataset

VI. CONCLUSION AND ONGOING WORKS

Proposed graph code used edge dictionary and adjacent edges information to preserve the structural information of the original graph. Instead of expensive pair-wise comparisons, it can be efficiently used to detect automorphic graphs. Instead of path or subgraph decomposition process which could result in structural information lost and exhausted enumeration time, GCTrie is used to support all query types. From our experimental results, proposed approach outperforms the existing methods in index construction time and subgraph query response time. Similarity query processing is going to be observed as our ongoing work.

ACKNOWLEDGEMENT

This research was supported by University of Computer Studies, Yangon and University of Information Technology. We thank our colleagues who provided insight and expertise that greatly assisted the research.

REFERENCES

- [1] Y. W. Hlaing and K. M. Oo, "A graph representative structure for detecting automorphic graphs," in the Proceeding of 9th International Conference on Genetic and Evolutionary Computing, August 2015, pp.189-197.
- [2] S. Sakr and G. Al-Naymat, "Graph indexing and querying : a review," in International Journal of Web Information Systems, vol. 6 No.2, 2010, pp.101-120.
- [3] Y. W. Hlaing and K. M. Oo, "Graph code based isomorphism query on graph data," in the Proceeding of 2015 IEEE International Conference on Smart City/SocialCom/SustainCom together with DataCom 2015 and SC2 2015 (SmartCity 2015), Dec 2015, in press.
- [4] D. Shasha, J. T. L. Wang, R. Giugno, "Algorithmic and Applications if Tree and Graph Searching", 2002.

- [5] OrientDB <https://en.wikipedia.org/wiki/OrientDB>
- [6] R. N. Chittimoori, L. B. Holder, and D. J. Cook, "Applying the SUBDUE substructure discovery system to the chemical toxicity domain," in Proceeding of the 12th international Florida AI, Research Society Conference, 2003, pp.90-94.
- [7] D. A. Piriya Kumar, and P. Levi, "An Efficient A* based algorithm for optimal graph matching applied to computer vision," in GRWSIA-98, Munich, 1998.
- [8] D. Dupplaw, and P. H. Lewis, "Content-based image retrieval with scale-spaced object trees," in Proceeding of SPIE: Storage and Retrieval for Media Databases, Volume 3972, 2000, pp.253-261.
- [9] Adjacency List http://en.wikipedia.org/wiki/Adjacency_list.
- [10] Adjacency Matrix http://en.wikipedia.org/wiki/Adjacency_matrix.

AUTHORS PROFILE

Y. W. Hlaing received her bachelor degree in Computer Science (2006) from university of Computer Studies, Yangon and her master degree in Computer Science (2008) from Computer University, Mawlamyine. She is now a Ph.D candidate at University of Computer Studies, Yangon.

K. M. Oo received her B.Sc. in Mathematics from Yangon University; M.I.Sc. from Institute of Computer Science and Technology(ICST); and Ph.D. in Information Science from University of Computer Studies, Yangon(UCSY), Myanmar in 1994, 1996, and 2007 respectively. Currently, she is an associate professor in the Faculty of Computing, University of Information Technology (UIT), Myanmar. Her research interests include graph theory, and data mining.

An Efficient Migration Scheduling Algorithm Based on Energy Consumption and Deadline in Cloud Computing

Sogol Davati

Department of Computer Engineering
South Tehran Branch, Islamic Azad University
Tehran, Iran

Ahmad Khademzadeh

Education and International Scientific Cooperation
Department
Iran Telecommunication Research Center (ITRC)
Tehran, Iran

Abstract— Cloud is one of the most important shareable environments which several customers are connected to it in order to access services and products. In this environment all the resources are available in an integrated environment and several users can simultaneously send their requests. In this condition some methods are required to schedule in an efficient way. Important condition is when a cloud server is overloaded or request of user has deadline constraint. In these situations to provide efficient service, migration occurred. Migration of virtual machines can be used as a tool for achieving balance. In other words when load of virtual machines in a physical machine increased, some of the virtual machines migrate to another physical machine so that virtual machines do not face any shortage of capacity. Virtual machine migration is used for many reasons like reducing power consumption, load balancing in virtual machines, online maintenance and capability of responding to requests with deadline constraint.

In this thesis, we present an efficient algorithm for scheduling live migration of virtual machines which have applications with deadline constraint. This scheduling uses thresholds for migrating virtual machines in order to guarantee service level agreement (SLA). Simulations show that the proposed method acts better in number of received requests with success and level of energy consumption.

Keywords- Cloud computing; Virtual machine migration; Threshold; Cloudsim; SLA; Scheduling

I. INTRODUCTION

With new and innovative ideas for new internet services, software developers don't need to spend a lot of their investments on preparation of hardware and hiring people for triggering their services. They don't need to worry about overprovision of un-popular services or under provision of popular services. This feature is considered as a major advantage for information technology corporations. Because instead of preparation of servers and software infrastructures they have ability to focus on innovation and construction of business values.

Cloud computing is an instance of distributed computing in large scale for economies of scales that is managed like a pool of computing and storage power, platform and virtual services. Cloud computing has dynamic scalability which is delivered to remote costumers in internet on demand.

Cloud computing provides infrastructure, platform and software for costumers as a subscription-based services that are available on demand. These services are called infrastructure as a service, platform as a service and software as a service respectively. In other words, cloud computing is a model for supporting everything as a service.

Virtualization technologies change the method of datacenter in resource consumption of servers. Instead of using particular servers for each type of applications, virtualization provides the ability of observing resources as a pool of integrated resources. Therefore complexity reduces and management ability is much easier. By virtualization we can hide operation details from users. One of the most important virtualization techniques is virtual machines migration issue, i.e., transfer the allocated process to another virtual machine or cloud in order to release load and respond cloud requests effectively.

In fact, virtualization in cloud provides the ability of separating the applications and services from physical hardware. Therefore with virtualization, several virtual machines run on a single physical machine concurrently and use its capacity (processing, memory and network). Cloud provider should obey service level agreement and avoid its violation. Agreement in this level can include items such as respond time, job loss rate, energy consumption level, deadline, number of migrations and etc. In fact we should achieve a tradeoff between SLA parameters. It should be noted that because cloud is really huge and also load of virtual machines are not predictable, we can't avoid SLA violation completely.

The SLA that we focus on in this thesis is the combination of considering deadline constraint and reducing energy consumption. There are several studies about virtual machine migration.

II. RELATED WORKS

The main idea of live migration algorithm was proposed by Clark et al [1] in 2005. First of all, hypervisor marks all the pages as a dirty page then, algorithm transfer dirty pages over the network iteratively while the number of remain pages for transfer become less than threshold or the maximum iteration

achieved. After that hypervisor marks transferred pages as a clean page.

In 2006, Khanna et al [2] study about union and integration of servers. In this method they define thresholds for efficiency of resources to avoid application performance degradation. This method doesn't consider the network topology on migration algorithm details.

In Bobraff method in 2007 [3], number of requests calculated in regular intervals. Then virtual machines mapped on physical machines.

In 2007 wood et al [4] used virtual machine migration for balance establishment between resource requests in physical hosts but minimizing cost of servers was not their goal. This method uses grey box and black box approach to reduce critical zone in virtual clusters by using all facilities of virtual machine live migration.

Bradford et al in 2007 [5], proposed a system to support clear and live migration of virtual machines that use local storage in their stable states. This method is clear for migrated virtual machines and doesn't create pause in network open connections and this guarantee the compatibility of virtual machine in source and destination after migration.

In 2008, Tal Maoz et al [6] present an approach for migrating virtual machines and processes in grid and cluster environments. Requirements such as high service accessibility, resource consumption improvement and management increase the need of virtual machine migration.

In 2009, Van et al [7] proposed architecture for management of virtual machines in clouds based on migration scheduling. This method has long execution time.

In 2010, Jing Thai Piao et al [8] present a virtual machine placement and migration approach to minimize the time of data transfer. Simulation results show that proposed method is effective in optimization of data transfer between virtual machines so it helps the optimization of general application performance.

In 2011, RicaSinha et al [9] present an algorithm for scheduling live migration based on energy consumption that uses dynamic threshold. In this method dynamic threshold uses to increase processor efficiency in datacenter.

In 2011, Aziz Murtazaev et al [10] proposed a method that first of all sorts nodes based on load of virtual machine in decreasing order. Then virtual machines in last node of the list (node with lowest load) candidate for migration, virtual machines also sort based on weight in decreasing order. After that virtual machines one by one allocated to first node (node with highest load) and if effort was not successful, then the second node chooses and so on. Because the node with lowest load usually has the lowest number of virtual machines, so number of migration decreases.

In 2012, Ahmed Abba et al [11], proposed a method for scheduling deadline based on jobs in grid environment. This method doesn't consider the migration. Proposed algorithm chooses the job with nearest deadline for execution. This method uses dynamic time quantum. Jobs sort based on their

delay in increasing order. Then the job with minimum delay candidate for execution. If delay of two jobs were equal, then job candidate based on entrance time. Proposed algorithm compare with Round Robin and FCFS scheduling algorithm based on average execution time, average waiting time and maximum delay. Results show the improvement of these parameters.

In 2013 [12], the method was proposed for virtual machine scheduling in cloud datacenter based on energy efficiency. Results show that this algorithm only decreases energy consumption.

Quentin Perret et al in 2013 [13] proposed a scheduler based on deadline for distributed systems to discover solution for data placement management. This algorithm called cloud least laxity first and is compared with timeshared and space shared scheduling algorithms. In this method each job laxity defined as difference between deadline and remaining time to complete the job. This is a non-preemptive scheduling because in multiprocessing algorithm one of the assumptions is that if one job stops in one processor it can continue in another processor without considering its cost. Since information saves locally here, so if a job transfer to another node must execute from the beginning.

In 2013, Baghshahi et al [14] proposed a virtual machine migration method based on intelligent agents. In this paper, multiple virtual machines are considered as a cluster. These clusters are migrated from a data center to another data center with using weighted fair queuing.

III. VIRTUAL MACHINE LIVE MIGRATION

Live virtual machine migration is a technique that migrates the entire OS and its associated application from one physical machine to another. The Virtual machines are migrated lively without disrupting the application running on it. The benefits of virtual machine migration include conservation of physical server energy, load balancing among the physical servers and failure tolerance in case of sudden failure. Live migration is an extremely powerful tool for cluster and cloud administrator. An administrator can migrate OS instances with application so that the machine can be freed for maintenance. Similarly, to improve manageability, OS instances may be rearranged across machines to relieve the load on overloaded hosts. To perform the live migration of a virtual machine, its runtime state must be transferred from the source to the destination while virtual machine still running.

There are two major methods: Post-Copy and Pre-Copy memory migration. In the Post-copy method first suspends the migrating Virtual machine at the source node then copies minimal processor state to the target host and resumes the virtual machine, and begins fetching memory pages over the network from the source node.

There are two phases in Pre-copy method: Warm-up phase and Stop-and-Copy phase. In warm up virtual machine memory migration phase, the hypervisor copies all the memory pages from source to destination while the virtual machine is still running on the source. If some memory pages change during memory copy, they will be re-copied until the rate of recopied

pages is not less than page dirtying rate. In Stop and Copy phase, the VM will be stopped in source and the remaining dirty pages will be copied to the destination and virtual machine will be resumed in destination. [15]

IV. PROPOSED METHOD

In this method we use Pre-Copy live virtual machine migration. First of all we define number of virtual machines, physical machine and applications (jobs). There is an assumption that each virtual machine only can respond to one request (task) in a moment and other requests placed in a queue. Resources (CPU, network, memory, I/O) are allocated to each virtual machine in order to handle workload and ensure SLA. Another assumption is that the placements of applications are definitely clear from the beginning and resources are allocated to them. Based on the agreement between customer and cloud service provider, each job has a specific deadline for execution and if it doesn't execute in specific time SLA violation occurred. So to prevent SLA violation, virtual machines that execute the job must migrate.

Generally the proposed method present in four levels:

- diagnose suitable time for virtual machine migration
- select suitable virtual machine for migration
- define migration time sequence for selected virtual machine
- select suitable destination for migrated virtual machine

A. Diagnose suitable time for virtual machine migration

Hosts check periodically and in addition to energy consumption below specific threshold and load balancing. Another goal is to find hosts that can respond to received requests in defined deadline, i.e., sum of waiting time and execution time of all their virtual machines is greater than sum of defined deadline. So that best time for virtual machines migration is when their hosts consumption are higher than specific threshold or host has virtual machine that it can't respond to received request in defined deadline. It isn't necessary to migrate all virtual machines of candidate hosts. In each level only one virtual machine will migrate. Then virtual machines of each candidate host will placed in migration queue. Fig.1 shows the migratable virtual machines queue in each selected host.

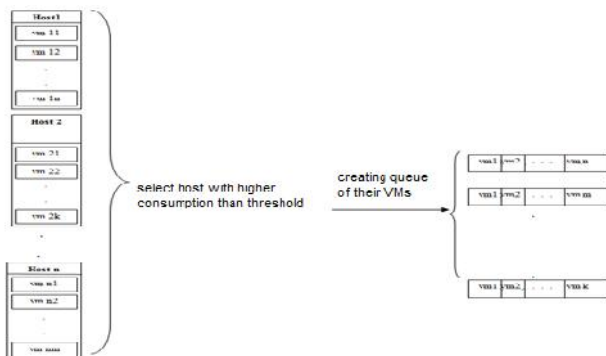


Figure 1. migratable Virtual machines queue

In fact in this level the policy is to find which host should migrate its virtual machine. Equation (1) shows the calculation of virtual machine's waiting time which is the time interval between virtual machine creation and job execution beginning time.

$$VmWaitingTime = JobStartExeTime - VmCreationTime \quad (1)$$

Equation (2) shows virtual machine execution time estimation. Job execution time is depend on number of executable statement in one second and number of its processing elements (Processor)

$$VmExecutionTime = \frac{JobLength}{VmMIPS * VmNoPEs} \quad (2)$$

Equation (3) shows the calculation of requested job deadline. Deadline is a function of execution time and has many ways to calculate. Defined deadline shouldn't set very low because it leads to more SLA violation and it also shouldn't set very high because in this situation we don't face any SLA violation at all. Therefore here the considered job one time executes on virtual machine with highest speed and next time execute on virtual machine with lowest speed. Based on following relation we set a weight for each execution time and $k=0.6$ ensures the considered balance.

$$Deadline = BestCase + K * (WorstCase - BestCase) \quad (k=0.1, \dots, 0.9) \quad (3)$$

After calculation of above relations, host that sum of execution and waiting time of its virtual machines is greater than sum of deadline of virtual machine's considered host or energy consumption of the host is greater than defined threshold, because of avoiding SLA violation and quality of service degradation that is based on consumption level and deadline, can candidate as a host with high energy consumption:

$$\begin{aligned} A &= VmExecutionTime + VmWaitingTime > Deadline \\ B &= Utilization > UtilizationThreshold \\ \text{if} [(A)OR(B)] &\rightarrow Hostsselected \end{aligned} \quad (4)$$

1. input.
2. Hosts one by one
3. virtual machine list of each host
4. Current Time
5. Utilization Threshold
6. Virtual machine Selection Policy
7. Calculate Vms Waiting Time
8. Calculate Vms Execution Time
9. Calculate Vms Best Execution Time
10. Calculate Vms Worst Execution Time
11. Calculate Vms Deadline
12. for (virtual machines of a Host)
13. if (Utilization > UtilizationThreshold) OR (VmWaitingTime + VmExecutionTime > Deadline)
14. Host added to OverUtilized Host list
15. Sort Virtual machine List of OverUtilized Hosts in increasing order of Deadline

Figure 2. pseudo code of proposed scenario

B. Select suitable virtual machine for migration

In this level based on proposed scenario in each selected physical host, virtual machines put in a queue. We select a virtual machine that its host can't respond to its request in defined deadline. For each virtual machine of selected host following method will execute:

First of all, by (1), virtual machine waiting time will be calculated. Then by (2) virtual machine execution time will be calculated. Here we consider the remaining of execution time i.e., the spent time of requested job subtracted from total execution time. Finally by (3), deadline will be calculated. Here, we consider the remaining time of deadline i.e., the spend time of deadline subtracted from total deadline. After these calculations virtual machine that sum of its execution time and waiting time is greater than deadline is candidate for migration because of Sla violation and QOS reduction avoidance.

$$Sum = VmExecutionTime + VmWaitingTime$$

$$if (Sum > Deadline) \rightarrow Migration\ Required \quad (5)$$

C. Define migration time sequence for selected virtual machine

In this level, migration sequence of selected virtual machines will be defined. After finishing the first level, the virtual machine queue id created and is sorted based on increasing deadline.

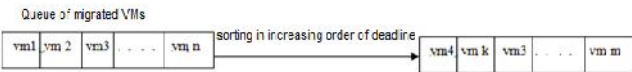


Figure 3. An example of defining the sequence for VM migration

D. Select suitable destination for migrated virtual machine

In this level, every host will be checked. For host that after accepting of migrated virtual machine based on defined scenario doesn't place in high consumption hosts group, following calculations take place:

$$VmMigrationTime = \frac{VmRam(MB)}{\frac{HostBW(Gb/s)}{2}} \quad (6)$$

The main part of a virtual machine that will be migrated is its memory which size and amount of that will affect the migration time. So the time of migration is based on amount of memory (information) that host can send in a second. Bandwidth divided by 2 because the half of bandwidth will be used for migration and the other half will be used for communication with destination host.

Now virtual machine migration time determine in hosts respectively in order to find suitable destination for virtual machine. Since some part of the job execute before and during migration, the remaining length of the job to use in execution time calculation achieve from (7):

$$RemainLength = jobLength - (VmMIPS * (CurrentTime - JobStartExTime)) \quad (7)$$

In this level The remaining time of execution will be calculated based on (2) and then the remaining Time of deadline will be calculated based on (3). After that (8) will be checked:

$$If (VmRemainExeTime + VmMigrationTime < RemainDeadline) \quad (8)$$

After checking the above condition we will check the migration of virtual machine to a host in order to have lowest energy consumption during migration. Amount of energy consumption during migration achieve from (9) [14]:

$$PowerDiff = PowerAfterAllocation - HostPower$$

$$PowerAfterAllocation = \frac{HostUtilizationMIPS + VmRequestedMIPS}{HostTotalMIPS} \quad (9)$$

TABLE I. DESCRIPTION OF USED PARAMETERS

| Parameters | Description |
|----------------------|--|
| VmWaitingTime | Time that takes until execution of VM starts |
| JobStartExeTime | Start time of job execution in VM |
| VmCreationTime | Time of VM creation in host |
| VmExecutionTime | Time that takes until VM execution finished |
| JobLength | Job length of VM based on million instructions per second. In fact this parameter shows the number of instructions that should execute to complete a job |
| VmMIPS | Number of million instructions that execute in a second |
| VmNoPEs | Number of VM's processing elements |
| Deadline | Deadline of VM in order to execute its job |
| BestCase | best-case of requested job execution time |
| WorstCase | worst-case of requested job execution time |
| RemainLength | remain length of job |
| VmRemainExeTime | Vm remaining execution time |
| RemainDeadline | remaining time of deadline |
| VmMigrationTime | Vm migration time |
| VmRam | Ram of migrated VM |
| HostBW | Bandwidth of host |
| PowerDiff | Difference between host energy consumption after migration and energy consumption of host |
| PowerAfterAllocation | Host energy consumption after migration (VM allocation) |
| HostPower | Host power consumption |

| Parameters | Description |
|---------------------|--|
| HostUtilizationMIPS | Host utilization before migration in million instructions per second |
| VmRequestedMIPS | Migrated Vm requested million instruction per second |
| HostTotalMIPS | Total million instructions that host can execute in a second |

Generally, proposed algorithm will execute again after a specific time slot (5min) in order to check hosts condition. In this level execution and deadline remaining time will be evaluated. Reasons for using proposed method to schedule virtual machine migration:

- Load balancing
- Decrease energy consumption caused by virtual machine migration and disable idle servers.
- Suitable for environments that have to respond to requests in specific time and therefore ensure the SLA between customer and cloud service provider

Afterward, we will check implementation and evaluation the results.

```

1. input :
2. HostList
3. vm that needs migration
4. Current Time
5. initialize minPower as max value of double
6. for eachHost do
7. Calculate Vm remain Execution Time
8. Calculate Vm remain Deadline
9. Calculate Vm Migration Time
10. Calculate sum of Exe Time and Migration Time
11. Calculate Power after allocation
12. If(sum <= Deadline)
13. find host that migration to it has minimum power consumption
14. add host to allocated host

```

Figure 4. pseudo code of Select suitable destination for migrated virtual machine

V. IMPLEMENTATION OF PROPOSED METHOD

For implementation of proposed algorithm we use cloudsim simulator that use java programming language in eclipse environment.

A. Definition of cloudsim

There are several simulators for cloud computing simulation that most of them are based on java. Cloudsim simulator, support system components behavior modeling like datacenter, virtual machines and preparing policies. Components like datacenter, resources, broker and cloud information service are defined as an entity in cloudsim.

Support of virtual machine migration is a cause of using this simulator.

B. Simulation of proposed method in cloudsim

In this simulation, we use one cloud and migration is between physical machines in this cloud. In this scenario we use one datacenter which has following features:

- System architecture : x86
- Operating system : Linux
- Virtual machine monitor : XEN

In this datacenter, there are 50 physical machines with following features:

- Number of physical machine types : 2 types
- Number of processing elements : 2
- Amount of ram : {4096,8192} which is based on type of physical machine
- Amount of storage : 1GB
- Amount of bandwidth : 1GB/s

Type of physical machines calculates as follows in each stage:

$$\text{hostType} = i \% \text{No of Host Type} \quad (10)$$

Features of virtual machines are as follows:

- Number of virtual machine types : 10 types
- Number of million instructions per second : between 500 and 2500 MIPS which is based on type of virtual machine
- Number of processing elements : 1
- Amount of bandwidth : between 100 and 118 Mb/s
- Size of virtual machine : 2.5 GB

Type of virtual machines calculates as follows in each stage:

$$\text{vmType} = \frac{i}{\frac{\text{vmsNumber}}{\text{No of vm Type}}} \quad (11)$$

Features of requested job :

Here, we consider 6 different types for jobs that each type has different start time and length. Type of requested job calculates as follow in each stage:

$$\text{CloudletType} = \frac{i}{\frac{\text{cloudletsNumber}}{\text{No of Cloudlet Type}}} \quad (12)$$

C. Comparing proposed method with existing methods based on energy consumption

Till now many methods proposed for decreasing energy consumption in cloud computing environment. One of the most important methods, is the one that proposed by cloudsim simulator development group. Beloglazov et al in 2012 [16], proposed a method for dynamic integration of virtual machines based on virtual machine resource consumption data analyze. This method caused decrease in energy consumption with service level agreement guarantee. This thesis uses combination of several methods for selection of virtual machine for migration policy and virtual machine allocation policy. Using the combination of local regression minimum migration cause decrease in energy consumption and better results. This selection policy uses Loess method that proposed by Cleveland, for selection of host that it's consumption is greater than threshold.

For virtual machine selection, this method selects the virtual machine that has lowest migration time. Migration time estimated based on amount of consumed ram by virtual machine divided by amount of host available bandwidth. Results of this method compare with proposed algorithm based on energy consumption, number of virtual machine migration and number of host's shutdown. Existing and proposed algorithm executed in eclipse environment and cloudsim simulator. Number of physical and virtual machines differs in each experiment.

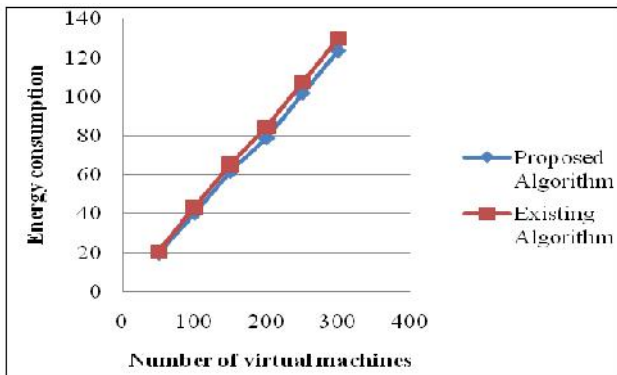


Figure 5. Comparison of energy consumption

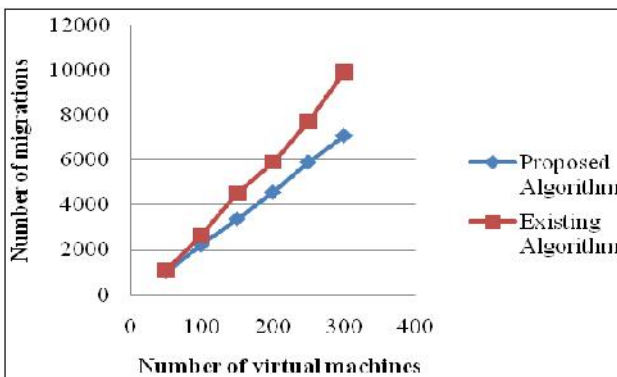


Figure 6. Comparison of number of migration

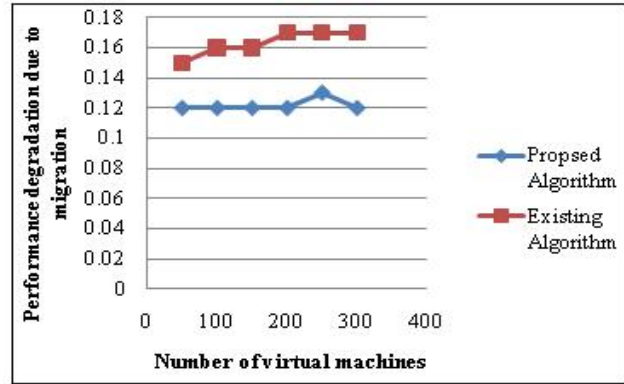


Figure 7. Comparison of performance degradation due to migration

Since in proposed method, virtual machines that sum of their execution time and waiting time is greater than their defined deadline candidate for migration from the beginning, number of unnecessary migrations decrease and because the migration process causes overhead on system, in proposed method energy consumption decreases. Decrease in energy consumption with decrease in number of host's shutdown shows that migration is efficient and has good performance. As seen in diagram 3 performance degradation due to virtual machine migration to another host and ratio of unplaced virtual machines to all migration requests is much lower in proposed method. One of the reasons is reduced number of host's shutdown because this function causes performance degradation and system overhead.

D. Comparison of proposed and existing methods based on deadline

Many thesis and researches use deadline in scheduling which most of them didn't use migration. One of this thesis proposed in 2014 by hadadian et al [17] that uses scheduling based on predefined deadline and balance factor in cloud computing environment and results compare with FIFO and Round Robin algorithms. The goal of this method is to achieve minimum cost and reduce unsuccessful jobs and therefore better performance. In this method received jobs sort based on their deadline then the best cloud chooses for job execution and algorithm again executes in specific intervals in order to make changes if necessary. This method increases success ratio. Finally the cloud that has maximum balance factor and it's response time consider the deadline, selects as a suitable cloud. This method doesn't consider virtual machine migration and is proposed as a future works to improve load balancing.

This simulation uses 6 different data for execution time and deadline. In each of them, there are 10 jobs (tasks) which each of them have specific execution time and deadline. We compare our proposed method with this method based on minimum, maximum and average rate of request loss.

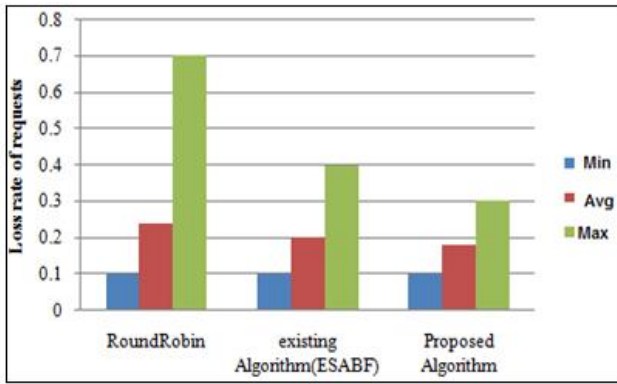


Figure 8. Comparison of loss rate

This compare show that average and maximum loss rate in proposed method have been improved, i.e. proposed method can complete more requests successfully. Proposed method considers energy consumption too and uses migration when response time in one host becomes greater than deadline and for reducing energy consumption. Therefore if the host can't response the received request, virtual machine migration to another host with ability to response on time occurred and because of migration, request's loss rate will be reduced.

VI. CONCLUSION

As mentioned earlier, internet and IT world that are vital part of human life are growing day by day. Needs of community members like information security, fast processing, online and dynamic access, power of focusing on organization's projects instead of keeping servers and saving costs are very important too. Nowadays solution in technology is cloud computing. Cloud computing is service delivery instead of product and it uses virtualization technology by internet.

In this thesis we provide request's response within their defined deadline and consider energy consumption as an important parameter. If the host can't response or energy consumption becomes greater than threshold, virtual machine migration to another host with ability to response virtual machine after migration occurred. Energy consumption must be lower than threshold after migration. This algorithm will be iterated in specific intervals so in every stage of simulation, status of virtual machine could be checked and migration will happen if necessary. Although this method increases the computation overhead, evaluation and comparison of results show that this method is much better in energy consumption in compare with existing methods and with reducing unnecessary migrations it causes reducing in migration overheads. So level of reducing performance and cost of migration is lower in compare with existing methods. Proposed method can response better in defined deadline in compare with presented method in [16] and average and maximum loss rate of requests are lower than existing method.

REFERENCES

- [1] C.Clark, K.Fraser, S.Hand, J.Gorm ,E.Jul, C.Limpach, I.Pratt, and A.Warfield," Live Migration of Virtual Machines" , NSDI '05: 2nd Symposium on Networked Systems Design & Implementation, 2005.
- [2] G.Khanna,K.Beaty,G.Kar,A.Kochut,"ApplicationPerformance Management in Virtualized Server Environments", 2006
- [3] N.Bobroff,A.Kochut,and K.Beaty,"Dynamic Placement of Virtual Machines for Managing SLA Violations", 2007.
- [4] Wood,"Black-box and Gray-box Strategies for Virtual Machine Migration", 2007.
- [5] R.Bradford,E.Kotsovinos,A.Feldmann,and H.Schi"oberg, " Live Wide-Area Migration of Virtual Machines Including Local Persistent State", 2007
- [6] T.Maoz,A.Barak,and L.Amar, and H.Diwanji,"Combining Virtual Machine Migration with Process Migration for HPC on Multi-Clusters and Grids", 2008.
- [7] N.Van,and D.Tran," SLA-aware virtual resource management for cloud infrastructures", October 2009.
- [8] J.Tai Piao,and J.Yan,"A Network-aware Virtual Machine Placement and Migration Approach in Cloud Computing", Ninth International Conference on Grid and Cloud Computing, 2010.
- [9] R.Sinha, N.Purohit," Power Aware Live Migration for Data Centers in Cloud using Dynamic Threshold ", 2011.
- [10] A.Murtazaev,and S.Oh,"Sercon: Server Consolidation Algorithm using live Migration of Virtual Machines for Green Computing", May2011.
- [11] A.Abba, B.Zakaria,and S.N.M Shah,"Design, Development and Performance Analysis of Deadline Based Priority Heuristic for Job Scheduling on a Grid", International Conference on Advances Science and Contemporary Engineering, 2012.
- [12] Ch.Ghribi, M.Hadji,and D.Zeghlache"Energy Efficient VM Scheduling for Cloud Data Centers: Exact allocation and migration algorithms", 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing, 2013.
- [13] Q.Perret,G.Charlemagne,S.Sotiriadis,andN.Bessis,"A Deadline Scheduler for Jobs in Distributed Systems", 27th International Conference on Advanced Information Networking and Applications Workshops, 2013.
- [14] A.Beloglazov,and R.Buyya,"Optimal Online Deterministic Algorithms and Adaptive Heuristics for Energy and Performance Efficient Dynamic Consolidation of Virtual Machines in Cloud Data Centers", Concurrency And Computation: Practice And Experience, 2012
- [15] L.Baghshahi,A.Khademzadeh,and S.jabbehdari, "Migration of Virtual Clusters with Using Weighted Fair Queuing Method in Cloud Computing", International Journal of Computer Science and Information Security, Vol. 11, No. 12, December 2013.
- [16] D.Patel,M.Karamta, M. D. Bhavsar, andM. B. Potdar, "Live Virtual Machine Migration Techniques in Cloud Computing: A Survey", International Journal of Computer Applications,January2014
- [17] M.Hadadian,and A.M.Rahmani,"TasksScheduling Algorithm with Predefined DeadLine and Considering the Balance Factor", International journal of Computer Science & Network Solutions, Aug.2014

EEG Discrimination of Rats under Different Architectural Environments using ANNs

Mohamed.I.El-Gohary
Department of Physics, Al-Azhar University,
Cairo, Egypt

Tamer.A.A.Al-Zohairy
Department of Computer science, Al-Azhar University,
Cairo, Egypt

Amir. M.M. Eissa
Department of Physics ,Al-Azhar University,
Cairo, Egypt

Sally .M. Eldeghaidy
Department of Physics ,Suez canal University,
Suez, Egypt

Hussein.M.H.Abd El-Hafez
Department of Physics ,Al-Azhar University
Cairo, Egypt

Abstract— The present work introduces a new method for discriminating electroencephalogram (EEG) power spectra of rat's brain housed in different architectural shapes. The ability of neural networks for discrimination is used to describe the effect of different environments on brain activity. In this research the rats were divided into four groups according to the type of environmental shapes as: control (normal cage), pyramidal, inverted pyramidal and circular. The brain activities (EEG) were recorded from rats of each group. Fast Fourier Transform (FFT) analysis of EEG signals was carried out to obtain power spectra. Two different neural networks are used as classifiers for power spectra of the different 4 groups: multi-layer perceptron (MLP) with backpropagation and radial basis function RBF networks with unsupervised K means clustering algorithm. Experimental studies have shown that the proposed algorithms give good results when applied and tested on the four groups. The multilayer with backpropagation and radial basis function networks achieved a performance rate reaching 94.4 % and 96.67% respectively.

Keywords: EEG, Architectural shape, artificial neural networks, power spectrum, Backpropagation Algorithm, Radial basis function network.

I. INTRODUCTION

Electroencephalography (EEG) is the recording of electrical potential difference during brain activity. EEG signals are often used for observation and diagnosis of neurological disorders [1, 2]. EEG provides information about the functional state of the brain more than structural information.

Environment is the combination of external physical conditions that surround and influence the growth, development and the survival of organisms. Due to the consideration of architectural environment that surrounds the human being; many trends appear in architecture aiming to achieve environmental balance and human comfort [3].

Since the energy is the ability to produce an effect, the present scientific view of reality supports the idea that we are composed of energy fields and the wave is the principle shape of energy. All the levels of energy react with each other through resonance to produce some kinds of harmony or energy balance. This affects the energy balance of human and living organisms in the form of waves or radiation energy that have specific wavelength and frequency. The human himself produces electromagnetic fields that enable him to react with any other objects in the surrounding environment that have their own frequency or vibration [4]. So we could say that two similar geometrical shapes have similar frequency, vibrations and motion energy that transfer into the human by resonance. The architectural design of a room may create a new medium that will have an influence on the physical and physiological state of the bodies [5].

Research on Pyramids proof some evidence that space within the great pyramid and its small replicas intensifies and/or generates the energy of electromagnetic radiations and other forms of the so-called universal energy [6]. The effect of this 'pyramidal energy' had been studied on solids [7], liquids [8], plants [9], microorganisms [10], animals [11] furthermore actually human volunteers. Some of the discoveries of such studies include rapid growth of plants, faster healing of bruises and burns, longer preservation of milk [12], and an enhanced vitalization and better relaxation in human subjects [13]. A Various number of volunteers have expressed that meditating inside the pyramid was easier than meditating outside as they felt more peaceful, more relaxed and less distracted [14].

Some Previous studies were carried out to show the effect of architectural shapes on rats, exposure to pyramidal environment reduced neuroendocrine and oxidative stress and increased antioxidant defense in rat [15], provided better

wound healing and protection against stress-induced neurodegenerative changes [16]. EEG recorded from rats housed into different shapes showed an enhancement in alpha brain waves for rats housed in pyramidal shape compared with other shapes [17]. In this research we will use artificial neural networks to discriminate between EEG of rat's brain housed in different architectural shapes.

Artificial Neural Networks (ANNs) are designed to perform information processing like a human brain. ANN is a mathematical model which performs calculating, decision making and learning. It is based on a human nervous system; therefore, ANN components simulate neural networks. In an ANN, processing element, weight, activation function and output correspond respectively to cell body, synapse, dendrite and axon in a biological neural network of the brain [18].

Artificial neural networks (ANNs) are widely used in science and technology with applications in various branches of chemistry, physics, and biology [19]. ANNs have proven useful in the analysis of blood and urine samples of diabetic patients [20], leukemia classification [21] and EEG recording for prediction of epileptic [22]. In addition, ANNs have also been applied in the diagnosis of colorectal cancer [23], colon cancer [24] and early diabetes [25].

In the present research, we use two different architectures of supervised artificial neural networks, namely multi-layer perceptron (MLP) with backpropagation and radial basis function RBF networks to discriminate EEG power spectra from the four different environmental shapes.

This paper is organized as follows: in section II, materials and methods are described. The two types of neural networks, as a classifier, are described in section III. In section IV, the results and discussion of the study are provided. Finally, conclusions are remarked in section V.

MATERIALS AND METHODS

Forty male adult albino rats greater than 150 g were used in the present study. Rats were kept in an animal house under constant laboratory conditions, fed and provided water for 3 weeks. The biological clock of the animals was kept as normal.

A. Architectural Shapes of Rats Housing

As shown in figure 1, the animals were divided into 4 groups (10 for each):

- The first group was considered as a control in a wired cage.
- The Pyramidal shape has dimensions of 21.5 cm height, 50 cm base and 22 cm for each side. The four triangular sides

of the pyramidal shape are angled upwards at an angle of 51° with the base.

- The inverted pyramidal shape has dimensions of 25 cm height, 83 cm base and 50 cm side. The four triangular sides of the inverted pyramidal shape are angled upwards at an angle of 129° with the base.
- Circular shape has dimensions of 28 cm height, 55 cm diameter.

All the latter 3 geometrical shapes are made of wood with a wire mesh door at the top [26].

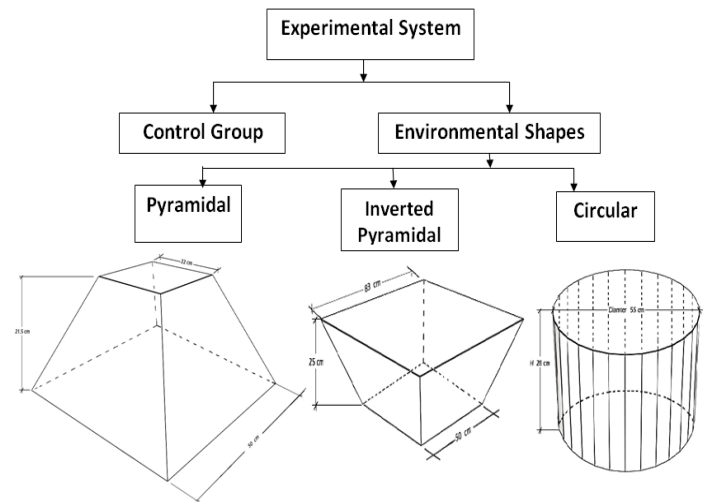


Figure 1. The classification of different experimental groups

B. EEG Recording Method

Rats were anesthetized with sodium pentobarbital (50 mg/kg, intra-peritoneal) prior to electrode implantation. Four epidural permanent stainless electrodes were implanted in the skull in positions overlying both motor and visual cortices of rat's brain. A fifth electrode (reference electrode) was implanted in the contra lateral crest of the skull (nasal) at the interaural line [27]. The five implanted electrodes were connected to the amplifier and then to the analogue to digital converter (A/D).

As shown in Fig.2 The EEG signal from the rat passes through electrodes to jack box, amplifiers, filters, analogue to digital converter and finally to the computer display. The recording was made several times in order to ensure that the EEG recordings are stable and the rat's brain is free from any lesions. After recording EEG signals from the rat's brain, FFT analysis was carried out for EEG signals to obtain power spectra.

The spectra resulting through this analysis are divided into four frequency bands: Delta (δ) (1-4 Hz), Theta (θ) (4-8 Hz), Alpha (α) (8-13 Hz) and Beta (β) (14-30 Hz) [28, 29]. A calculation of the mean power of amplitude (μV^2) for EEG signals for each band was performed [29]. The recorded EEG

signals were grouped into four main categories based upon the environmental shape; (1) pyramidal, (2) inverted pyramidal, (3) circular and (4) control.

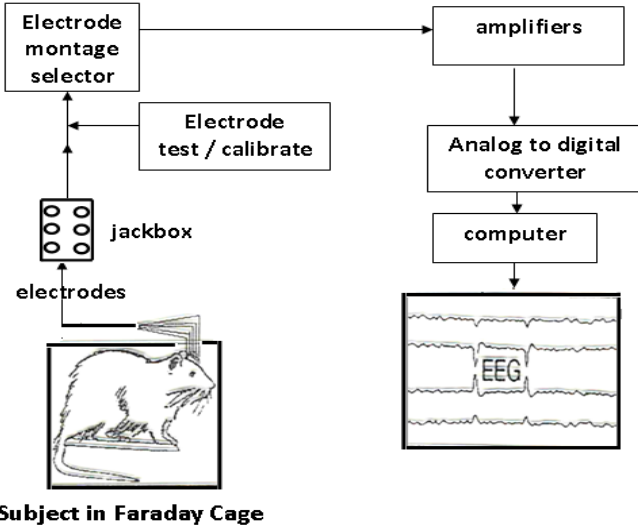


Figure 2. Block diagram of the main functional components of the EEG recording system.

C. Input data preparation

Each rat had 4 electrodes connected to it. This comprised 4 mono-polar EEG channels as well as 4 bipolar EEG channels. Data from each of the 8 channels was divided into four different frequency bands delta, theta, alpha and beta. So far we obtain 32 bands from each weekly recording. Data from 3 weeks presented as a vector of 96 data points. Forty samples from the four groups of length 96 data point are presented to the neural network for classification.

II. ARTIFICIAL NEURAL NETWORKS

Artificial Neural Networks (ANNs) are the interconnection of simple processing nodes which functionality is modeled from the neurons in the brain. The ANN consists of an input layer, an output layer and at least one hidden layer. The input values to the network are fed from the input layer passing through the hidden layer to the output layer. The input values are forwarded to the nodes in the hidden layer. The values obtained as inputs are processed within hidden layer and forwarded to either the nodes of the next hidden layer or nodes of the output layer.

The processing capacity of the network is determined by the relative weights of the connections in the network. Supervised learning is the process of changing the weights of the links between the nodes of the network to map patterns presented at the input layer to target values of the output layer. This is done using training algorithms.

In the following subsections two different types of neural networks are presented to be used for the classification of EEG signals from different architectural environments [30].

A. MLP with BP Algorithm Network

In this subsection the structure of the MLP neural network with the backpropagation (BP) algorithm used to adjust its parameters for classification purposes is introduced.

1) Multilayer Perceptron

The multilayer neural network architecture can have any number of layers. Fig. 3 illustrates network architecture with 4 layers, the first layer is the input layer and the last layer is the output layer. In between first and last layers are second and third layers which are the hidden layers. In this work two hidden layers are used to classify EEG of the 4 different shapes represented by features vector.

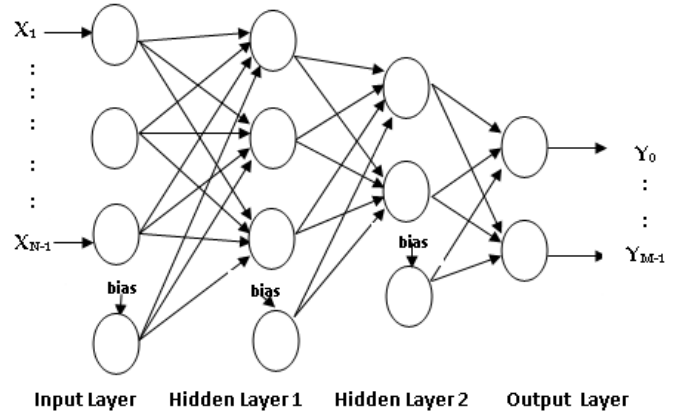


Figure 3. Multilayer perceptrons architecture.

2) Backpropagation Algorithm

Back propagation algorithm is the most popular in the supervised learning architectures. It is considered a generalization of the delta rule for nonlinear activation functions.

Learning of the MLP by Backpropagation algorithm has two phases. First, a training input pattern is presented to the input layer. The network propagates the input pattern from layer to layer until output pattern is generated by the output layer. If this pattern is different from the desired output, an error is calculated and then propagated backward through the network from the output layer to the input layer. The weights are modified as the error is propagated. The backpropagation training algorithm is an iterative gradient descent approach designed to minimize the mean square error between the actual output of multilayer feedforward perceptron and the desired output. It requires continuously differentiable non-linearity. We assume a sigmoid logistic nonlinearity as a transfer function of the network nodes.

Following is the summary of the BPN algorithm:

Step 1: Initialize weights and offsets.

Set all weights and node offsets to small random values.

Step 2: Present input and desired outputs.

Present a continuous valued input vector x_0, x_1, \dots, x_{N-1} and specify the desired output d_0, d_1, \dots, d_{M-1} . If the network is

used as a classifier then all desired outputs are typically set to zero except for that corresponding to the class the input is from. That desired output is +1. The input could be new on each trial or sample from a training set could be presented cyclically until stabilization.

Step 3: Calculate actual output.

First calculate the formulas

$$I_j = f_j^{h1} \left[\sum_i w_{ji}^{h1} x_i + b_j^{h1} \right] \quad (1)$$

$$Z_j = f_r^{h2} \left[\sum_j w_{rj}^{h2} I_j + b_r^{h2} \right] \quad (2)$$

Where x_i is the i^{th} network input, w_{ji}^{h1} is the connection weight from the i^{th} input to the j^{th} neuron in the 1st hidden layer, w_{rj}^{h2} is the connection weight from the j^{th} neuron in the 1st hidden layer to the r^{th} neuron in the 2nd hidden layer, b_j^{h1} is the weight from the bias to the j^{th} neuron in the 1st hidden layer, b_r^{h2} is the weight from the bias to the r^{th} neuron in the second hidden layer, $f_j^{h1}(\cdot)$ and $f_r^{h2}(\cdot)$ are nonlinear sigmoid activation functions defined as

$$f(\text{netinput}) = \frac{1}{1 + e^{-\text{netinput}}} \quad (3)$$

The network output is calculated by the following equation

$$\hat{y}(k+1) = f_k^O \left[\sum_k w_{kr}^O Z_r + b_k^O \right] \quad (4)$$

where w_{kr}^O is the weight connection of the k^{th} neuron in the output layer to the r^{th} neuron in the 2nd hidden layer, b_k^O is the bias weight for the k^{th} output neuron, and f_k^O is the transformation function between 2nd hidden layer and output layer which is a linear function.

Step 4: compute the error signal

$$e_j(k) = d_j(k) - \hat{y}_j(k) \quad (5)$$

Step 5: Adapt weights.

Use a recursive algorithm starting at the output nodes and working back to the first hidden layer. Adjust weights by

$$w_{ij}(k+1) = w_{ij}(k) + \eta \delta_j x_i \quad (6)$$

In this equation $w_{ji}(k)$ is the weight from hidden node i or from an input to node j at time k , x_i is either the output of node i or is an input, η is the learning rate, and δ_j is an error term for node j . If node j is an output node then

$$\delta_j = (d_j - \hat{y}_j) \quad (7)$$

where d_j is the desired output of node j and \hat{y}_j is the actual output.

If node j is an internal hidden node then

$$\delta_j = x_j(1 - x_j) \sum_k \delta_j w_{jk} \quad (8)$$

Where k is over all nodes in the layers above node j .

Internal node thresholds are adapted in a similar manner by assuming they are connection weights on links from auxiliary constant-valued inputs equal to 1.

Step 6: Repeat by going to step 2.

B. Radial Basis Function Neural network

In this subsection the structure of the RBF neural network and the algorithms used to adjust its parameters for classification purpose are introduced.

1) Structure of RBF neural Network

Fig.4 shows a typical RBF network, with q inputs (x_1, \dots, x_q), and P outputs (y_1, \dots, y_p). The hidden layer consists of h computing units connected to the output by h weight vectors ($\alpha_1, \dots, \alpha_h$).

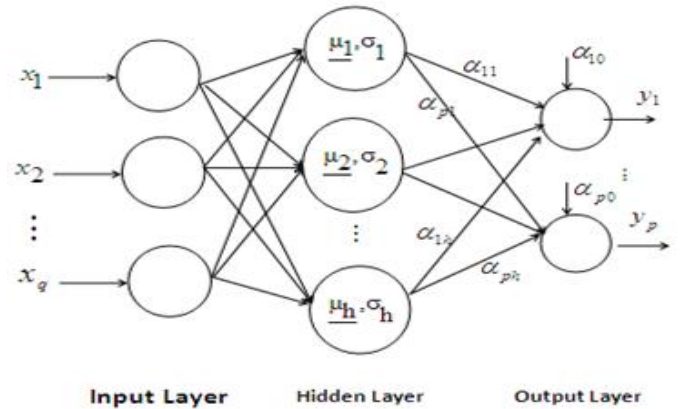


Figure 4. RBF Network architecture.

Response of one hidden unit to the network input at the i^{th} instant, x_i , can be expressed by

$$\Phi_k(x_i) = \exp \left[-\frac{1}{(\sigma_k^i)^2} \|x_i - \underline{\mu}_k^i\|^2 \right], \quad (k:1 \dots h) \quad (9)$$

where $\underline{\mu}_k^i$ is the center vector for k^{th} hidden unit at i^{th} instant, σ_k^i is the width of the Gaussian function at that time, and $\|\cdot\|$ denotes the Euclidean norm. The overall network response is given by

$$\hat{y}_i = f(x_i) = \underline{\alpha}_0^i + \sum_{k=1}^h \underline{\alpha}_k^i \Phi_k(x_i) \quad (10)$$

where $\underline{y}_i \in R^p$, $\underline{x}_i \in R^q$. The coefficient vector $\underline{\alpha}_k^i$ is the connecting weight vector of the k^{th} hidden unit to output layer, which is in the vector form of $\underline{\alpha}_k^i = [\alpha_{1k}^i, \dots, \alpha_{lk}^i, \dots, \alpha_{pk}^i]^T$. Thus the coefficient matrix of the network can be expressed as

$$A_{p \times h}^i = \begin{bmatrix} \alpha_{11}^i & \dots & \alpha_{1k}^i & \dots & \alpha_{1h}^i \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{l1}^i & \dots & \alpha_{lk}^i & \dots & \alpha_{lh}^i \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{p1}^i & \dots & \alpha_{pk}^i & \dots & \alpha_{ph}^i \end{bmatrix}$$

and the bias vector is $\underline{\alpha}_0^i = [\alpha_{i0}^1, \dots, \alpha_{i0}^l, \dots, \alpha_{i0}^p]^T$.

2) Training Radial Basis Function Network

Training RBF neural network consists of determining the location of centers and widths for the hidden layer and the weights of the output layer. It is trained using a two-phase approach: in the first phase, unsupervised learning occurs, which its main objective is to optimize the location of center and width. In the second phase, the output layer is trained in a supervised mode using the least mean-square (LMS) algorithm to adjust the weights so as to obtain the minimum mean square error at the output. The following are the three steps of the hybrid learning method for an RBF neural network, and they are discussed in more details in the next three subsections:

- i. Find the cluster centers of the radial basis function; use the k-means clustering algorithm (calculation of center).
- ii. Find the width of the radial basis function using P-nearest neighbor.
- iii. Find the weights; use LMS (weight estimation).

i. Calculation of Centers

To calculate the centers of the radial basis function we use the k-means clustering algorithm. The purpose of applying the k-means clustering algorithm is to find a set of clustered centers and partition the training data into subclasses. The center of each cluster is initialized to a randomly chosen input datum. Then each training datum is assigned to the cluster that is nearest to itself. After training data have been assigned to a new cluster unit, the new center of a cluster represents the average of the training data associated with that cluster unit.

When all the new centers have been calculated, the process is repeated until it converges. The recursive k-means algorithm is given as follows:

1) Choose a set of centers $\{\underline{\mu}_1, \underline{\mu}_2, \dots, \underline{\mu}_h\}$ arbitrarily and give the initial learning rate $\gamma(0)=1$.

2) Compute the minimum Euclidean distance

$$L_i(k) = \|\underline{x}(k) - \underline{\mu}_i(k-1)\| \quad i: 1 \dots h \quad (11)$$

$$r = \arg[\min L_i(k)]$$

3) Adjust the position of these centers as follows:

$$\begin{aligned} \underline{\mu}_i(k) &= \underline{\mu}_i(k-1) + \gamma(k)(\underline{x}(k) - \underline{\mu}_i(k-1)) \quad (i=r) \\ &= \underline{\mu}_i(k-1) \quad (i \neq r) \end{aligned} \quad (12)$$

4) $k = k+1$, $\gamma(k) = 0.998\gamma(k-1)$ and go to 2.

ii. Width Calculation

After the RBF centers have been found, the width is calculated. The width represents a measure of the spread of data associated with each node. Calculation of the width is usually done using the P-nearest neighbor algorithm. A number P is chosen and for each center, the P nearest centers are found. The root-mean squared distance between the current cluster and its P nearest neighbors is calculated, and this is the value chosen for σ . So, if the current cluster center is $\underline{\mu}_j$, the value of width is given by:

$$\sigma_j = \sqrt{\frac{1}{P} \sum_{i=1}^P (\underline{\mu}_j - \underline{\mu}_i)^2} \quad (13)$$

A typical value of P is 2, in which case σ is set to be the average distance from the two nearest neighboring cluster centers.

iii. Weight Estimation

Learning in the output layer is performed after calculation of the centers and widths of the RBF in the hidden layer. The objective is to minimize the error between the observed output and the desired one. It is commonly trained using the LMS algorithm [32] and is summarized as follows:

Training sample: Input signal vector = $\underline{\Phi}(k)$

Desired response = $\underline{d}(k)$

User-selected parameter: $0 < \eta < 1$

Initialization: Initialize the weights $\hat{\underline{w}}(0)$.

Computation: For $k = 1, 2, \dots$ Compute

$$\underline{e}(k) = \underline{d}(k) - \hat{\underline{w}}^T(k) \underline{\Phi}(k)$$

$$\hat{\underline{w}}(k+1) = \hat{\underline{w}}(k) + \eta \underline{\Phi}(k) \underline{e}(k)$$

C. Training and Testing of the Two Different Neural Networks

In this part, Multilayer perceptron with backpropagation and radial basis function with k means clustering algorithm are used to solve classification problem. The two different NNs are programmed using C++ programming language [33]. The input

layer for both neural networks consists of 96 source nodes as mentioned before. The Initial weights for the two different neural networks are taken randomly from the interval [0, 1]. The value of the learning rate for the two neural networks was chosen experimentally in the range of (0.001–0.9) as we will see in the next section (IV) [34–36]. Choosing the number of the hidden layer nodes for MLP and RBF is shown in the section (IV). The number of output layer nodes depends on shapes needed to be discriminated. In our investigation, the EEG power spectra of four different shapes are used as samples. The corresponding desired outputs in both MLP and RBF are (0,1,0,0),(1,0,0,0), (0,0,1,0), (0,0,0,1) for control, pyramidal, circular and inverted pyramidal group respectively. The number of samples is 220 EEG power spectra for all shapes. Forty samples have been used as training data for the neural networks while 180 samples have been used for testing the neural networks.

III. RESULTS AND DISCUSSION

Before applying the two models of ANN for the discrimination of 40 samples of EEG power spectra of the different shapes, the network needs to be trained to optimize the network performance. For MLP with BP, different learning rates are taken in the range (0.001–0.9). Figure 5 shows the effect of different learning rates on the performance of MLP with backpropagation. The learning rate of 0.01 gives the best performance for NN. The effect of the number of hidden neurons for MLP is presented in Table I. With a fixed number of iterations of 100000, 60 first layer hidden nodes and 30 second hidden layer nodes resulted in the best performance compared with other combinations of hidden nodes.

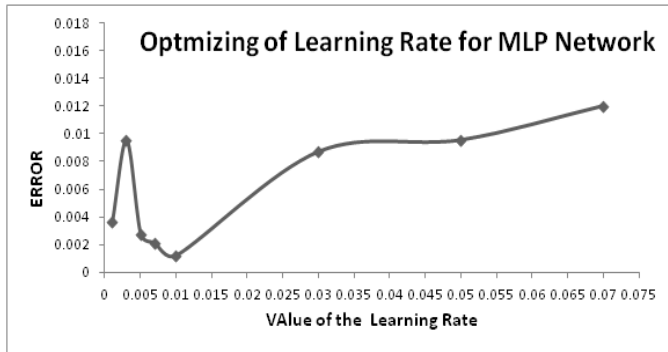


Figure 5. Influence of different Learning Rates on MLP neural network

For RBF network with k means clustering algorithm, Figure 6 shows the effect of different learning rates on the performance of network. The best value of the learning rate was 0.08. Table II shows the effect of the number of different RBF clusters with a fixed iteration number of 100000, the best performance was obtained when the number of clusters is 4. So a RBF network with 4 hidden layer nodes and the learning rate equal to 0.08 is used for testing.

TABLE I. EFFECT OF THE NUMBER OF HIDDEN NODES ON PERFORMANCE OF ANN (TESTING RESULTS)

| No of Nodes in 1st hidden layer | No of Nodes in 2nd hidden layer | Accuracy Training Control group % | Accuracy Training Pyramidal group % | Accuracy Training Inverted pyramidal group% | Accuracy Training Circular group% |
|---------------------------------|---------------------------------|-----------------------------------|-------------------------------------|---|-----------------------------------|
| 20 | 10 | 92.06 | 88.45 | 88.37 | 93.45 |
| 30 | 15 | 95.43 | 84.34 | 90.13 | 94.44 |
| 40 | 20 | 89.40 | 91.16 | 92.78 | 94.00 |
| 50 | 25 | 83.53 | 85.44 | 91.31 | 93.93 |
| 60 | 30 | 98.13 | 97.42 | 99.13 | 96.08 |
| 70 | 35 | 93.10 | 86.51 | 86.18 | 93.67 |

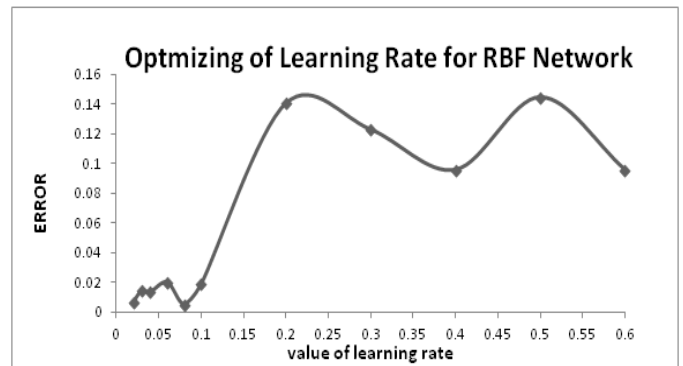


Figure 6. Effect of different Learning Rates on RBF neural network

TABLE II. EFFECT OF THE NUMBER OF CLUSTERS IN HIDDEN LAYER ON PERFORMANCE OF ANN

| No of clusters in hidden layer | Accuracy Control group % | Accuracy Pyramidal group % | Accuracy Inverted pyramidal group% | Accuracy Circular group% |
|--------------------------------|--------------------------|----------------------------|------------------------------------|--------------------------|
| 4 | 97.06 | 96.34 | 95.38 | 97.13 |
| 10 | 93.79 | 86.07 | 88.16 | 98.30 |
| 30 | 81.85 | 96.05 | 42.23 | 92.25 |
| 50 | 91.87 | 95.71 | 83.60 | 94.39 |
| 70 | 84.05 | 75.29 | 89.64 | 98.26 |

REFERENCES

After training both networks, selected test sets are used (50 from control group, 50 from pyramidal group, 40 from inverted pyramidal group and 40 from circular group). It is found that, the RBF neural network performed best with average classification accuracy of 96.67% with respect to the number of correctly classified instances (table IV), it was followed by MLP neural network with backpropagation algorithm with a classification accuracy of 94.44% as shown in Tables III.

TABLE III. PERFORMANCE OF MLP IN CLASSIFICATION OF EEG POWER SPECTRA FROM THE FOUR DIFFERENT SHAPES

| subjects | No of tested data | No of correct | No of incorrect | Correct % |
|--------------------|-------------------|---------------|-----------------|--------------|
| control | 50 | 47 | 3 | 94 |
| Pyramidal | 50 | 46 | 4 | 92 |
| circular | 40 | 39 | 1 | 97.5 |
| Inverted pyramidal | 40 | 38 | 2 | 95 |
| Total | 180 | 170 | 10 | 94.44 |

TABLE IV. PERFORMANCE OF RBF IN CLASSIFICATION OF EEG POWER SPECTRA FROM THE FOUR DIFFERENT SHAPES

| subjects | No of tested data | No of correct | No of incorrect | Correct % |
|--------------------|-------------------|---------------|-----------------|--------------|
| control | 50 | 49 | 1 | 98 |
| Pyramidal | 50 | 48 | 2 | 96 |
| circular | 40 | 39 | 1 | 97.5 |
| Inverted pyramidal | 40 | 38 | 2 | 95 |
| Total | 180 | 174 | 6 | 96.67 |

IV. CONCLUSION

In present study we have developed artificial neural networks for classification of EEG signals obtained from rats living in different architectural shapes. Two neural networks are used for the classification process which are RBF and MLP with backpropagation neural networks. A MLP with 60 nodes in the 1st hidden layer and 30 nodes in the 2nd hidden layer are used. The learning rate is given by $\eta=0.01$. RBF has 4 hidden layer nodes and the learning rate used is $\eta = 0.08$. From the test results using 180 EEG power spectra for environmental shapes, we found the best classification accuracy for RBF and MLP NNs are 96.67% and 94.46% respectively.

- [1] Agarwal, R., Gotman, J., Flanagan, D. and Rosenblatt, B., (1998). Automatic EEG analysis during long-term monitoring in the ICU. *Electroencephalography and clinical Neurophysiology*, 107(1), pp.44-58.
- [2] Batar H., (2007). *Analysis Of EEG Signals Using The Wavelet Transform And Artificial Neural Network*. M.Sc. Thesis, Süleyman Demirel University Graduate School of Applied and Natural Sciences, Isparta, Turkey.
- [3] Nahed, M., Salwa, A. L., Abdel Monsef, A., & Gehan, A. N., (2010). A study on radiation energy of Pyramidal shape 1-Effect of housing within a Pyramid model on cancer growth and some blood parameters of mice. *J. Rad. Res. Appl. Sci.*, Vol 3, No.4(A), pp.1211 - 1224
- [4] Richard and Iona M., (1989). *The holistic Qabalah. Psychedelic monographs Essays volume (4)*.
- [5] Nagy G.A. (2007). *The architectural figuration as a matrix for environmental control due to bioenergy science*. PhD thesis, Univ of Ain Shams, Egy: pp 67-93
- [6] Cazeau, C. J. (2012). *Exploring the unknown: Great mysteries reexamined*. Springer Science & Business Media.
- [7] Hook J.R., (2010). *Solid State Physics (2nd Edition)*. H.E. Hall, Manchester Physics Series, John Wiley.
- [8] Alekseev, Perepelkin I. (2002). Russian Patent No. 2184574. Russian patent office. Translation by <http://worldwide.espacenet.com>. Retrieved June 30, 2012
- [9] Abdelsamie, M. A., Rahman, R. A., & Mustafa, S. (2014). Pyramid Shape Power as a New Halal-Compliant Food Preservation and Packaging Technique. *Procedia-Social and Behavioral Sciences*, 121, 232-242.
- [10] Kumar I. R., Swamy NVC & Nagendra HR, (2005). Effect of pyramids on micro-Organism. *Indian J Traditional Knowledge*. 4 (4); 373–379.
- [11] DeSalvo, J. (2003). *The Complete Pyramid sourcebook*. 1st Book Library. 137.
- [12] Gopinath R K, Prem Anand Nagaraja & Nagendra H R, (2008). The effect of pyramids on preservation of milk, *Indian J Traditional Knowledge*. 7 (2) 233-236.
- [13] Heline, C. (1996). *Healing and regeneration through color. Color Healing: Chromotherapy*, 46.
- [14] Toth M., Nielsen G. (1985). *A falcon's eye view of the pyramid*. In: *Pyramid power*. USA: Inner Traditions India, pp: 17-23.
- [15] Bhat S., Rao G., Murthy KD. , Bhat PG. (2003). Effect of housing rats within a pyramid on stress parameters. *Indian J Exp Biol*, 41: 1289-93.
- [16] Kamath S., Rao S.G. , Murthy K.D. , Bairy K.L. & Bhat S. (2006). Enhanced wound contraction and epithelization period in steroid treated rats: Role of pyramid environment. *Indian J Exp Biol*. 44: 902-4.
- [17] EL Gohary M.I, EL-Sayed E.M, Nagy G.A, EL-Sayed T.M, Hussein H.M, (2013). Influences of Architectural Shapes of environment on the Electrical Activity of Rat's Brain. *Egy J of biophysics & Biom Eng*. Vol(11). 1-6 .
- [18] Dogali, G., & Bozkurt, M. R. (2015). The Detection of Normal and Epileptic EEG Signals using ANN Methods with Matlab-based GUI. *International Journal of Computer Applications*, 114(12).
- [19] Amato, F., Lopez, A., Peña-Méndez, E. M., Vañhara, P., Hampl, A., & Havel, J. (2013). Artificial neural networks in medical diagnosis. *Journal of applied biomedicine*, 11(2), 47-58.
- [20] Catalogna, M., Cohen, E., Fishman, S., Halpern, Z., Nevo, U., & Ben-Jacob, E. (2012). Artificial neural networks based controller for glucose monitoring during clamp test. *PLoS one*, 7(8), e44587.
- [21] Dey, P., Lamba, A., Kumari, S., & Marwaha, N. (2011). Application of an artificial neural network in the prognosis of chronic myeloid leukemia. *Analytical and quantitative cytology and histology/the International Academy of Cytology [and] American Society of Cytology*, 33(6), 335-339.
- [22] Fernandez-Blanco, E., Rivero, D., Rabuñal, J., Dorado, J., Pazos, A., & Munteanu, C. R. (2012). Automatic seizure detection based on star graph topological indices. *Journal of neuroscience methods*, 209(2), 410-419.

- [23] Spelt, L., Andersson, B., Nilsson, J., & Andersson, R. (2012). Prognostic models for outcome following liver resection for colorectal cancer metastases: A systematic review. *European Journal of Surgical Oncology (EJSO)*, 38(1), 16-24.
- [24] Ahmed, F. E. (2005). Artificial neural networks for diagnosis and survival prediction in colon cancer. *Molecular cancer*, 4(1), 29.
- [25] Shankaracharya, D. O., Samanta, S., & Vidyarthi, A. S. (2010). Computational intelligence in early diabetes diagnosis: a review. *The review of diabetic studies: RDS*, 7(4), 252.
- [26] Bhat, S., Rao, G., Murthy, K. D., & Bhat, P. G. (2003). Effect of housing rats within a pyramid on stress parameters. *Indian journal of experimental biology*, 41(11), 1289-1293.
- [27] Bringmann, A. (1995). Topographic mapping of the cortical EEG power in the unrestrained rat: peripheral effects of neuroactive drugs. *Archives italiennes de biologie*, 133(1), 1-16.
- [28] Campbell, I. G., & Feinberg, I. (2009). Longitudinal trajectories of non-rapid eye movement delta and theta EEG as indicators of adolescent brain maturation. *Proceedings of the National Academy of Sciences*, 106(13), 5177-5180.
- [29] Lüchinger, R., Michels, L., Martin, E., & Brandeis, D. (2012). Brain state regulation during normal development: Intrinsic activity fluctuations in simultaneous EEG-fMRI. *NeuroImage*, 60(2), 1426-1439.
- [30] Ehlers, C. L., Desikan, A., & Wills, D. N. (2013). Developmental differences in EEG and sleep responses to acute ethanol administration and its withdrawal (hangover) in adolescent and adult Wistar rats. *Alcohol*, 47(8), 601-610.
- [31] Alrajeh, K. M., & Alzohairy, T. A. (2012). Date fruits classification using MLP and RBF neural networks. *International Journal of Computer Applications*, 41(10), 36-41.
- [32] Lee, C. K. M., Ho, W., Ho, G. T., & Lau, H. C. (2011). Design and development of logistics workflow systems for demand management with RFID. *Expert systems with applications*, 38(5), 5428-5437.
- [33] Hossain, M., Rahman, M., Prodhan, U. K., & Khan, M. (2013). Implementation Of Back-Propagation Neural Network For Isolated Bangla Speech Recognition. arXiv preprint arXiv:1308.3785.
- [34] Rakotomamonjy, M. A., Migeon, B., & Marche, P. (1998). Automated neural network detection of wavelet preprocessed electrocardiogram late potentials. *Medical and Biological Engineering and Computing*, 36(3), 346-350.
- [35] Zurada, J. M., (1997) :Introduction to Artificial Neural Network Systems, West Publishing Company, St. Paul, MN.
- [36] Hassoun, H. M., (1998): Fundamentals of Artificial Neural Networks, Printice-Hall of India Private Limited, New Delhi.

Forward Error Correction for Storage Media: An Overview

Kilavo Hassan

Nelson Mandela African Institution
of Science and Technology, School
of Computational and
Communication Science and
Engineering, P.O. Box 447 Arusha,
Tanzania

Kisangiri Michael

Nelson Mandela African Institution
of Science and Technology, School
of Computational and
communication Science and
Engineering, P.O. Box 447 Arusha,
Tanzania

Salehe I. Mrutu

The University of Dodoma, College
of Informatics and Virtual
Education, P.O. Box 490 Dodoma

Abstract— As the adoption of Information and Communication Technology (ICT) tools in production and service rendering sectors increases, the demand for digital data storage with large storage capacity also increases. Higher storage media systems reliability and fault tolerance are among the key factors that the existing systems sometimes fail to meet and therefore, resulting into data loss. Forward error correction is one of the techniques applied to reduce the impact of data loss problem in digital data storage. This paper presents a survey conducted in different digital data storage companies in Dar es Salam, Tanzania. Data were collected and analyzed using Statistical Package for Social Sciences (SPSS). Secondary data were captured from user and manufacturer technical reports. It was revealed that data loss is still a predominant challenge in the digital data storage industry. Therefore, the study proposes the new storage media FEC model using locked convolutional encoder with the enhanced NTC-Viterbi decoder.

Index Terms—Storage Media, FEC, NTC, Viterbi, RS

I. INTRODUCTION

A storage media is any device on which data or information can be electronically stored, kept, and retrieved when needed [1, 2]. Storage media are devices that store user information (data) and application. The media can be categorized as optical data storage, magnetic hard disk drives, magnetic tape drives and flash disk drives[3]. The prevention of the record, information, and data stored electronically and the ability to retrieve them later or in future, require more than the safe keeping of the storage media [4]. There is little published scholarly work regarding the failure pattern of storage media and the key factors that affect their lifetime. Most of the information reported comes from storage manufacturers themselves. Since there is a huge amount of information stored and transferred among many storage media, data loss due to disk failure is a major issue that affects the reliability of the system. Reliability, fault tolerance and performance of storage media are the biggest concerns [3]. The demand for storage media increases everyday and it is estimated that over 90% of all the information and data produced in the world is being stored on magnetic media, and most of them are stored on hard disk drives [5]. On the other hand, data storage industry faces technological challenges due to the increase in demand and consistency. Currently, digital data have become one of the

most important parts nowadays and create the increasing demand of data storage systems[6]. The world has entered the information era and any area of life is immersed in information.

The demand for digital data storage with large storage capacity, higher reliability and fault tolerance, easier accessibility, better scalability and cheap management poses a remarkable challenge on the storage industry[7]. Essentially, the demand for data storage becomes more and more growing and data storage system needs to have high density, short access time, and fast input and output transfer rate[6]. International Data Corporation (IDC) shows that data are expanding at approximately 50% to 80% per year and in other places the growth rate is closer to 100% per year. Some of the catalysts for this growth include databases, Enterprise Resource Planning (ERP), Supply Chain Management, E-procurement, Content Management, Data Mining, Customer Relationship Management (CRM), Electronic Document Management (EDM), emails, social media and multimedia[8, 9]. Seemingly, the demand for bigger, and faster memories has led to significant improvement in the conventional memory technologies like hard disk drives and optical disks. However, there is a strong proof that these two dimensional storage technologies are approaching their fundamental limits[2]. Almost all vital data are now stored on external disk storage subsystems. An average usable capability is approximately 2.18 peta bytes which is up to 12.8% year after year. Factors such as growth in storage requirement, larger capacity disks and subsystems, and affordable pricing have led to larger storage configuration [3]. Microsoft has introduced the functionality of storage pools and storage spaces to Microsoft Windows server 2012 and Microsoft Windows 8. This shows that we still need to have an improvement in the storage media[10, 11]. Increasingly, it is not viable to talk about a storage media these days without talking about solid state or flash storage [12, 13]. The use of portable storage media has increased, thus, there is a need for improving the security mechanism for such devices. A secure solution that provides more security and user convenience is needed to avoid direct data loss and compromising the security of that data[14].

More or less, one and all who have suffered data loss know the importance of reliable data storage. The error on a storage

device might be random or burst. There are some common causes of error in the storage devices. These error sources could be a scratch on the disc, the error from read/write failure, dust, or controller [7, 15]. It was observed that some of the challenges include managing storage growth, lack of skilled cloud technology professionals, lack of knowledge about forward error correction in storage media, designing and managing backup recovery and archive solution, and finally convincing the management to adopt the cloud. Figure 1 shows the distribution of hard disk problems at the Internet archive. In Figure 1, storage media problems such as disk failure, disk error and other disk subsystem problems consist of 60% of all the recorded hardware problems[16]. The common problem is the disk failure which covers 42% of all problems.

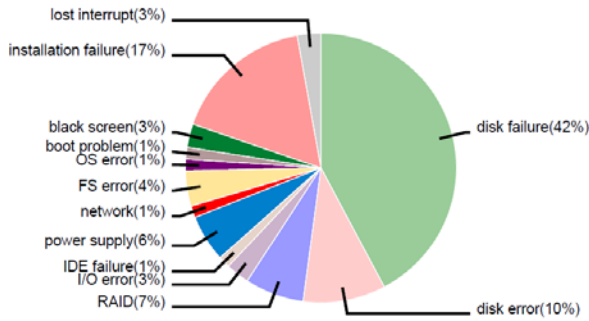


Figure 1: Hardware Failures Distribution at Internet Archive: Source : [16].

The recent research conducted in some companies in Dar es Salaam shows common error facing the storage media. In Figure 2, the recent study findings reveal that the most common error is the disk failure.

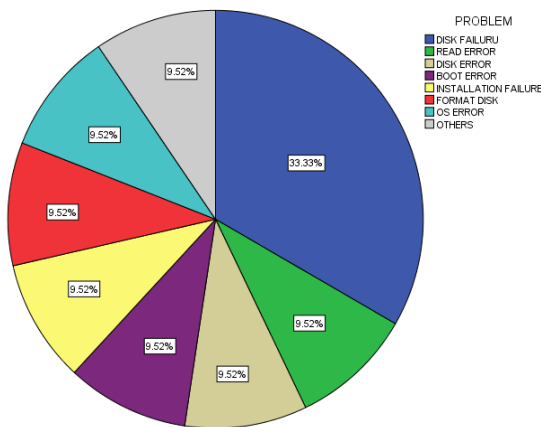


Figure 2: Common Storage Media Error for Different Companies
Reliability and performance of storage media are the biggest concerns. This implies that there is a need for the research to improve storage technologies in the existing systems.

II. STATE OF THE ARTS

In Spite of the problems discussed in the previous section, digital data storage industry applies different mechanisms such as backup and recovery systems to avoid data loss during failures. This section discusses efforts undertaken .

A: Backup System

Backup is the process of copying files or databases in order to preserve data in case of equipment failure or any other calamity. Backup or archiving data are later used to recover systems to their original state after failure. Backup has two main goals. The primary goal is to recover data after a loss by either deletion or corruption, whether intentionally or accidentally. The secondary goal is to restore the system to its earliest condition. There are three common types of backup systems which are full backup, incremental backup, and differential backup. In full backup, all the files are backed up every time you run a backup system. Incremental backup brings back the files that have changed since the last backup was done. Finally, differential backup provides backup of the files that have changed since a full backup was performed. However, different studies indicate that the pace of data growth is very high per year and even higher for the companies having large data and intensive application or distributed data centres. This means all data need to be backed up and duplicated. This fact makes the system more expensive and time consuming with this solution[17]. The crucial data should always be backed up to protect them from loss or corruption. Saving only one backup file sometimes might not be enough to preserve the information. In same vein, in order to safeguard crucial data, one needs to keep at least three copies of any such data, that is, one primary file and two backups. The two backup files are supposed to be kept in different media storage types to protect different types of hazards and making sure one file is kept offsite[18-20]. The importance of sensitive data to be backed up cannot be avoided because data storage media reliability can be low. With respect to this scenario, new techniques are introduced from time to time to secure vital data. For example, Oracle introduced sun ZFS appliance for Oracle Exadata Database which can restore up to 7TB/hour[21, 22]. Disaster planning would not be necessary if we were sure that nothing will ever happen in our environment such as hardware problem, viruses, and user errors. However, if a disaster is anticipated to occur, making backup and recovery systems is important to any vital data stored into the system. Essentially, the goal here is to create a way that crucial data will always be restored within acceptable time frame, within the budget, and without unnecessary shock on normal day to day activities[23]. However, one has to remember that it is very expensive because to back up data one needs extra devices and time to make backup, especially when there is large data.

B: Recovery System

Data recovery is a process of restoring or rescuing unintentionally deleted, corrupted, or inaccessible data to the storage media. Normally, this action is performed when we have physical damage or logical damage of the storage devices that prevent the file system of the devices to be mounted by the host system[24, 25]. Data recovery was introduced since 1974 when the magnetic data storage devices were commercially introduced into the market. And this was

after the users discovered the vulnerability of their data stored on those devices and quickly to overcome this problem data recovery was introduced to protect the data stored on those devices[8]. Any event that causes an interruption or destruction in the daily operation or processing for a period of time and affects that operation is called a disaster. This disaster should be addressed to recover the operation intended and this can be done by the disaster recovery which must ensure the continuation of all processes in the event when a disaster occurs [26]. In the world where we live today, tera bytes or peta bytes of data are not enough for storing large chunks of databases and therefore data recovery systems become more challenging especially in a distributed system[27]. The importance of a recovery system cannot be ignored and that is why even in the application such as Oracle you will find that the application for recovery system is included; for example, Oracle Recovery Manager (RMAN) utility from Oracle that was designed for online backup and recovery of Oracle database files[28]. Due to huge increase in the electronic data, large volume of storage devices is required to store those data. Currently, consumers prefer to store their data in cloud computing. However, if the cloud gets corrupted or damaged, then the consumers lose their important data. This led to the introduction of some mechanisms to back up data so that they could be restored at any time when the cloud fails. There is a technique like the plain data backup though it has many security and reliability problems. To overcome plain data backup and recovery problems, we can use a system like Redundant Array Independent Disk (RAID) [29]. Again we can recover data from Window search in which we obtain the record from search database either via carving or via extensible storage Engine API which provides a potential source of evidence about the files that we cannot access and the reason why it is not accessible[30]. The cost for keeping data safe could be taken into consideration because it has some financial implication. In this case, a company or an organization must decide on the best practices that will support the solution in such a way that it does not compromise the financial aspect[8, 9]. Data recovery is one of the fields whereby once you make a mistake it might lead you to irrecoverable data. So it is very important that you perform operations which you are familiar with and do not do any action which you are not familiar with. The importance of data recovery cannot be ignored as it is very significant for our data and that is why even in the Window system, the introduction of data recovery is always there. Remarkably, it is very expensive to have these systems, even in the case of softwares, some are very expensive. As long as people still lose their important data, something has to be worked on regardless the presence of these other technologies to recover the data when something happens to our systems.

C: FEC Associated

Forward Error Correction (FEC) is a digital signal processing technique used to improve data reliability[31]. FEC makes the

use of error correction codes to detect and correct errors automatically at the receiver. Error correction coding is the way whereby errors which may be introduced into digital data can be corrected based upon receiving data. Error detection coding is the means whereby errors can be detected based upon receiving the data. Collectively, error correction and error detection coding are error control coding[32, 33]. Forward Error Correcting (FEC) codes grant algorithms for encoding and decoding data bits, and help to accomplish data rates nearer to Shannon's Limit[34]. Similarly, error coding is a method of providing reliable digital data transmission and storage signal to noise ratio[35]. Error coding is used for fault tolerant computing in computer memory, magnetic and optical data storage media, satellite and deep space communications, network communications, cellular telephone networks, and almost any other forms of digital data communication[36, 37]. The design of efficient error correcting codes needs a complete understanding of the error mechanisms and error distinctiveness [38]. There are several types of codes. The first major classifications are linear and non linear. A linear code is an error correction code for which any linear combination of codes is also a codeword. Linear codes are partitioned into block codes and convolutional codes. Linear codes allow for more efficient encoding and decoding algorithm than other codes [39].

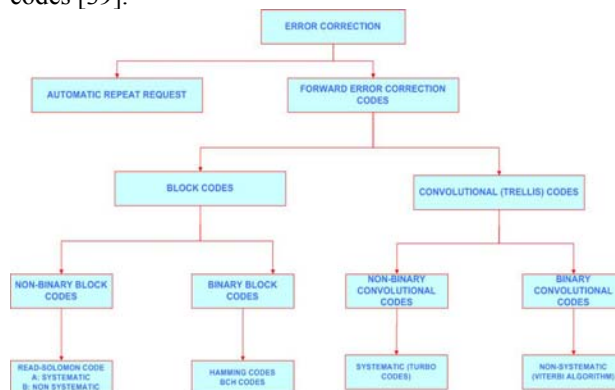


Figure 3: Error Correction Block

The linear codes are encoded using the method of linear algebra and polynomial arithmetics. If C is a linear code that a vector space over the field F has dimension k , then we say that C is an (n,k) linear code over F , or an (n,k) code, in short. Since linear codes allow for more and efficient encoding and decoding algorithm, then we will focus on linear codes and these are block codes and convolutional codes. Increasingly, a convolutional code operates on streams of data bits continuously, inserting redundant bits used to detect and correct errors. Convolutional codes are extensively used for real time error correction[40]. On the other hand, Block codes data are encoded in discrete blocks but not continuously. Convolutional codes are processed on a bit by bit basis. They are particularly suitable for the implementation in hardware, and the Viterbi decoder is one of the best algorithms in convolutional codes that allow optimal decoding. Additionally, Block codes are processed on a block by

block basis and the best algorithm is the Reed Solomon codes [41]. There are different algorithms which are used for error correction in the storage media. Currently, most of the error correction in the storage media is done by block codes and this is because block codes have the power of correcting burst errors. The convolutional codes are more powerful than the block codes besides there more computational complexity than block codes. Mrutu and his colleagues showed that Viterbi algorithm decoder has less computational complexity than other convolutional decoders[42]. Furthermore, the introduction of non transmittable code words technique to assist Viterbi decoders has enabled it to overcome bust errors [42]. This fact motivated the researcher to find out the effectiveness of the technique in storage media.

Reed Solomon.

Reed Solomon (RS) is an error correcting code that addresses multiple error correction especially burst errors in the storage media (hard disk drives, CD, and DVD), wireless and mobile communication units, satellite links, and digital communication. Reed Solomon is among the best block coding techniques of which the data stream to be transmitted is broken down into blocks and redundant data is then added to each block. The size of the block and the amount of check data added to each block is either specified for a particular application or can be user defined for closed system[43, 44]. RS codes are non binary cyclic codes with symbols made up of m bit sequences, where m is any positive integer having values greater than 2. Reed Solomon are block codes which are represented as RS (n,k) where n is the size of the code word length and k is the number of data symbols. The size of the code word length n is given by $2t + k$ where $2t$ is the purity symbol.

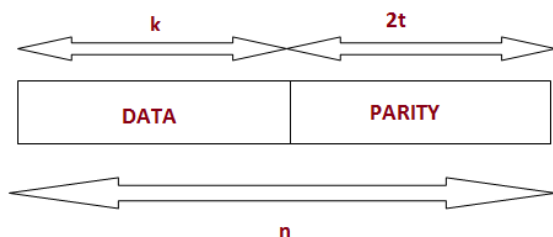


Figure 4: Structure of Reed Solomon

K =data symbol

$2t$ =parity symbol

N =symbol for RS code

Reed Solomon (RS) code was founded by Irving S. Reed and Gustave Solomon who by then were staff members at MIT Lincoln Laboratory. RS t was subsequently presented to the world on paper in 1960[45]. The RS code construction is based on the Galois Field ($GF(2^W)$) operation for W is positive integer. To encode Reed Solomon, we split the polynomial that representing the message by an irreducible generator polynomial, then the remainder is the RS code which we attach to the original message. The first commercial application for RS appeared in 1982 in compact disk (CD).

This was the first storage media device that used these error correction codes mechanisms [46]. RS illustrates a systematic way of building codes that may well detect and correct multiple random errors for different applications and they have strong burst and erasure error correction capabilities [47]. Also RS codes are able to correct multiple errors up to t errors, and they can be extended to correct errors up to $2t$ errors, just making sure that up to t errors the positions are known [48]. RS codes make them especially suited to the applications where burst error occurs and when there is a single bit error. This is because it does not matter to the code how many bits in a symbol are incorrect, if multiple bits in a symbol are corrupted, it only counts as a single error. This means for single bit errors, RS code is usually a poor choice. Seemingly, RS operates in two sides, in the encoder and in the decoder. The code operates on 8 bits and the number of symbols in the encoder block is $n=2^m-1$. The generator polynomial generates redundant symbols and they are appended to the message symbol. The decoder side is the one which identifies the location and magnitude of the error. The same generator polynomial is used to do the identification and the correction is applied to the received code.

Fundamentally, RS is suitable for much application because it has very high coding rate and low complexity. Besides being suitable for storage media, we still face problems in our storage media, and this alerts the researchers that there is a need to work more in improving the error correction mechanisms in the storage media. The Reed Solomon coding currently remains the optimal code for the smallest storage systems[49]. Generally, Reed Solomon is considered not very scalable[7]. Altogether, Reed-Solomon coding employs the same methodology. There are n data words and m element column vector. The product is an n data words + m element column vector representing the coding words[45, 50].

Convolutional Codes

Convolutional code is one of the error correcting codes which generate parity symbols through the sliding application of Boolean polynomial function $G(z)$ to the data stream. The sliding application is the representation of the convolution and its nature facilitates trellis decoding. More often Convolutional codes are termed as continuous. However, they have arbitrary block length, rather than being continuous, since most of the encoding is performed on blocks of data. Convolutional codes were introduced in 1955 by Peter Elias[51-53]. The main challenge in the convolutional codes is to find out the method for constructing codes of a given rate and minimizing their complexity[54]. Convolutional codes belong to FEC which makes use of the algorithm to automatically detect and correct error. They are one of the powerful error correcting codes that have a lower code rate (k/n). The common modulation scheme which is used is BPSK and QPSK, and the redundant bits are used to determine the error. The convolutional codes have two parts; one is the convolutional encoder and the second is the convolutional decoder. In the encoder, it uses encoder parameters which are n , k and m codes. Parameter ' k ' is input, ' n ' is output, and ' m ' is memory. Usually $k < n$ and ' m ' must be

large to achieve low error probability. The value of these encoder parameters can range from 1 to 8 for k and n, and from 2 to 10 for m[53, 55]. The decoder part is very important because the performance of the convolutional code is determined by the decoding algorithm and distance properties of the code, and that is why it is very important to determine the best decoding algorithm. For example, for small value of k, the Viterbi algorithm is commonly used. Viterbi decoder is one the most widely used decoding algorithm[53]

III: Methodology

It is through a combination of literature review, technical reports, and data obtained from selected employees in data storage sections through interviews, in some companies operating in Tanzania that we could assess the awareness of data storage techniques and error correction mechanisms. The primary data for this study were collected from digital data storage companies in Dar es Salaam, Tanzania. Secondary data were also collected from published papers and industrial technical reports. In this study, seven companies were surveyed to identify the kind of storage media they use, the reliability and failure rate, the main course of error or failure of storage devices, the possible measures to be taken when an error or failure occurs, the possible solution being suggested to avoid data loss when an error or failure occurs, and the future plan to avoid error in the storage media. The collected data were analyzed using Statistical Package for the Social Sciences (SPSS).

IV. RESULT AND THE PROPOSED MODEL

Analysis of reliability and failure behavior of the storage media from different companies from Dar es Salaam was conducted. All possible measures taken when a disk with error fails and the suggestion for the improvement for solving the problem were collected and analysed. When conducting a survey, the discussion with IT professionals working in the data storage departments was conducted. It was observed that most of the existing data storage media in the existing systems use Read Solomon [43, 44, 56]. However, the survey shows that data loss is still a problem in the industry. Therefore, there is a need of improving the algorithm to minimize the residual error during data reading process from digital storage media. It was also perceived that most users of digital storage media are not aware of FEC technologies in the media. In brief, the following are the observations from the literature review and data collected from different companies in Dar es Salaam.

- There is little scholarly published work on the error and failure pattern of the storage media and the key factors that affect their lifetime. Most of the information comes from storage manufacturers
- In spite of having RS codes for Forward Error Correction in storage media, the industry is still facing problems of data and mostly disk error and disk failure.
- Most of the error correction in the storage media is done by using RS codes which work best on Optical disk drive.

- RS is relevant in correcting burst errors and not appropriate in correcting single error because the errors are corrected in blocks so it is the wastage of resources.
- Convolutional codes are more powerful than Block codes; however, they are not preferred as a solution due to their computational complexity which causes results of data processing to delay.
- The introduction of Locked convolutional codes with Non Transmittable Codewords (NTC) at the decoding side can solve the problem of computational complexity in Viterbi decoder.
- Majority of the people are not aware and are not interested in improving the Forward Error Correction codes for storage media; rather they are interested in improving the backup systems and data recovery software.
- The hard disk drive is one of the storage media, which is mostly used to store data.

The following results in figures 5, 6, 7 and 8 are obtained from the data collected in seven companies in Dar es Salaam.

Figure 5 displays the reliability of storage media in some companies in Tanzania. The reliability ranges from daily to weekly basis in most of them. This demonstrates that the reliability is very low and that is why the backup is done on daily and weekly basis and very few of them do it yearly.

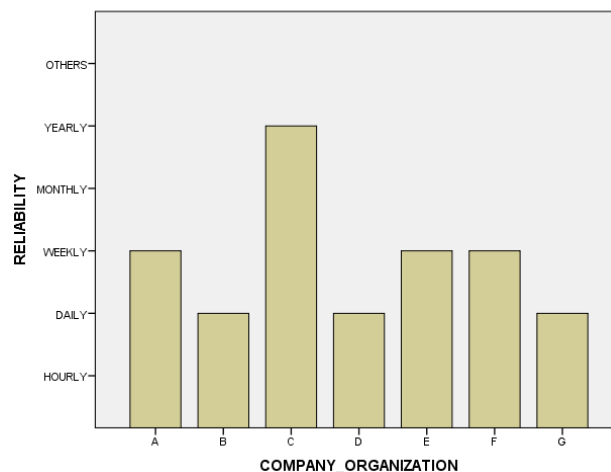


Figure 5: Reliability of storage media for some company in Tanzania

Figure 6 shows the failure rate of storage media in the surveyed companies. The failure rate is indicated for five years in the majority of the companies. This response to the findings indicated the reliability. If the reliability is low, then the failure rate is high. 85% revealed that the failure rate is not more than five years and only 14% is showed more than five years.

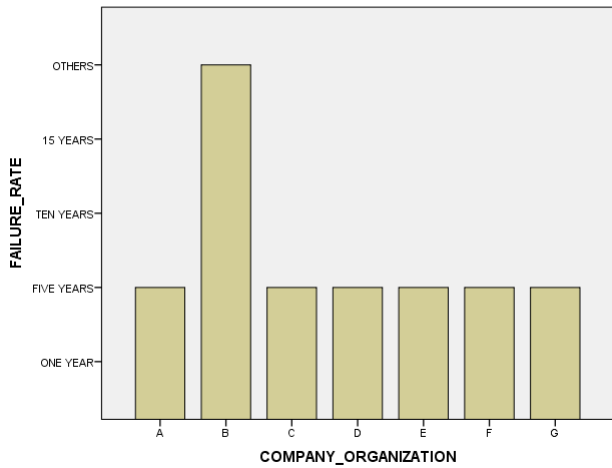


Figure 6: Failure Rate of Storage Media for Some Company in Tanzania
 Figure 7 illustrates that most IT professionals in the industry think that the best way to improve data storage system is through backup system and few through data recovery and restore system. It was revealed that 71.43% prefer to improve backup system, 14.29% prefer to improve on data recovery software and improve other techniques respectively. But 0% prefers the improvement of Forward Error Correction. The findings imply that most people are not aware of the importance of improving these algorithms. However, improving these algorithms can improve the reliability of the storage media. In this case, if deploying a backup system is very expensive then having a reliable storage system can help a lot, especially for those who do not store large volumes of data. Similarly, the individual case will also be improved as we all know that this issue can go down to individual level.

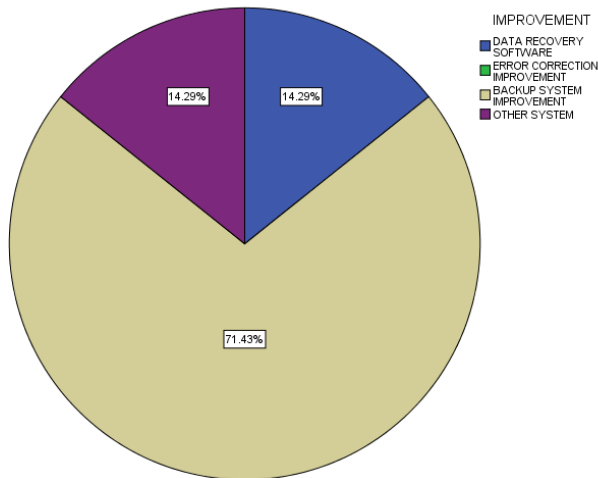


Figure 7: How to improve data storage loss

Figure 8 demonstrates that in most cases when there is data loss, very often the data are successfully recovered, and this is through backup system and data recovery software. But the analysis revealed that not always all the data are recovered. It is only 14.29% who are sure that the data will always be recovered and 85.71% remarked that most of the time they can recover their data but not always. 0% never lost everything

when they tried to recover the lost data, but the question is for those who do not have a backup system; thus it goes to an individual level.

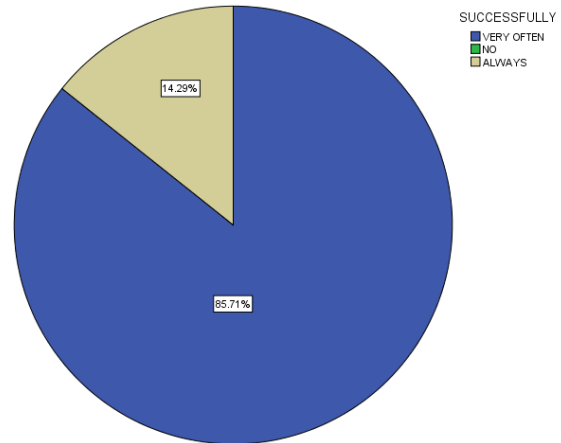


Figure 8: Data recovery after failure

Proposed Model

This study proposes improvement in the Forward Error Correction (FEC) in the storage media using Locked Convolutional Encoder and the enhanced Viterbi Algorithm with Non transmitted code words. In writing, the source data will be modulated, then encoded through locked convolutional encoder and at last written to the storage media. During reading the data will be decoded through Non transmitted code words (NTC)-Viterbi decoder then demodulated and finally the data can be retrieved. It is over two decades since Iterative Detection Read Channel (IDRC) technology was adapted in hard drive design[57]. This approach is delivers enhancement to signal and noise ratio that are unique and improves the reliability, resiliency, and overall storage capacity of the storage drives. The proposed model adopts the mentioned IDRC architecture as shown in Figure 8. This architecture has three levels of coding where modulation and locked convolutional codes are concatenated in the media writing process. A reverse concatenation is done on the reading process. A study on the design and evaluation of the proposed architecture is our next task. Error correction in magnetic field started in 1960s, with the application of Fire codes. Later on Reed Solomon (RS) took the major role in this area. RS is currently the principal error correcting codes being used in optical disc memories. RS stretched out when Compact Disc (CD) was introduced and the CD ROM were introduced at the beginning of 1980s [58]. However, data loss problem due to storage media failure is still a challenge in the industry, leaving a door open for researchers to find a more relevant solution

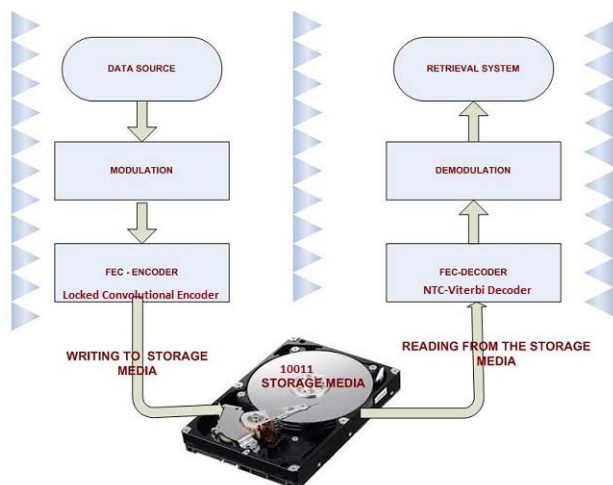


Figure 9: Proposed model

IV: CONCLUSION

This paper reviewed the challenges facing the industry on storage media. It has been revealed that data reliability can be increased by adopting or introducing powerful Forward Error Correction Code (FEC) in digital storage media. It has been witnessed that storage industries are confronted with disasters of varying degrees; yet, people in the industries are not interested in improving the storage devices reliability and fault tolerance, but are interested in improving the backup and recovery systems which have cost implication. Observation on introducing powerful FEC suggested the use of new technique which uses locked convolutional encoder and NTC-Viterbi decoder. The technique shows that if improved and adapted, may give better results compared to the current techniques. Therefore we recommend a further study on improving the new technique so that we can improve the current situation facing the storage industries.

REFERENCES

1. Martin, M.P. and R. Pileggi, *A quantitative analysis of Propolis: a promising new storage media following avulsion*. Dental traumatology, 2004. **20**(2): p. 85-89.
2. Goradia, I., J. Doshi, and K. Deulkar, *A Review Paper on 3D Optical Data Storage*. 2014.
3. Report, E., *MANAGING STORAGE: TRENDS, CHALLENGES, AND OPTIONS (2013-2014)*. 2014.
4. Van Bogart, J.W., *Magnetic Tape Storage and Handling: A Guide for Libraries and Archives*. 1995: ERIC.
5. Pinheiro, E., W.-D. Weber, and L.A. Barroso. *Failure Trends in a Large Disk Drive Population*. in *FAST*. 2007.
6. Nguyen, D.-C.L., Jaejin, *Performance Improvement Using Iterative Two-Dimensional Soft Output Viterbi Algorithm Associated with Noise Filter for Holographic Data Storage Systems*. 2014. **39**(3): p. 121-126.

7. Puttarak, N., *Coding for storage: disk arrays, flash memory, and distributed storage networks*. 2011.
8. Mayer, P., *Data Recovery: Choosing the Right Technologies*. Datalink White Paper, 2003.
9. Rao, Y.N., *DATA RECOVERY: CHOOSING THE RIGHT TECHNOLOGIES*. International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE), 2014. **11**(4): p. 27-29.
10. Liu, J., et al. *PLC-cache: Endurable SSD cache for deduplication-based primary storage*. in *Mass Storage Systems and Technologies (MSST), 2014 30th Symposium on*. 2014. IEEE.
11. Yoshiyama, M. and K. Toriyama, *Fujitsu's Approach to Server System Development*. FUJITSU Sci. Tech. J, 2011. **47**(2): p. 123-129.
12. Peters, M., *EMC's Comprehensive Flash Strategy*. 2014.
13. Yaakobi, E., et al., *Error correction coding for flash memories*. Flash Memory Summit, 2009. **160**.
14. Agwankar, N., et al., *Security For Portable Data Storage media*. 2014.
15. Tsai, T., N. Theera-Ampornpunt, and S. Bagchi. *A study of soft error consequences in hard disk drives*. in *Dependable Systems and Networks (DSN), 2012 42nd Annual IEEE/IFIP International Conference on*. 2012. IEEE.
16. Xin, Q., *Understanding and coping with failures in large-scale storage systems*. 2007, Technical Report UCSC-SSRC-07-06.
17. Xu, Q., et al., *YuruBackup: a space-efficient and highly scalable incremental backup system in the cloud*. International Journal of Parallel Programming, 2013. **43**(3): p. 316-338.
18. Crighton, I.P., *Data backup system*. 2001, Google Patents.
19. Gold, S., et al., *Data backup and recovery systems*. 2004, Google Patents.
20. Jewell, T.R., *Data backup system and method*. 2008, Google Patents.
21. Clarke, J., *Backup and Recovery*, in *Oracle Exadata Recipes*. 2013, Springer. p. 189-238.
22. Bach, M., et al., *Recovering Exadata*, in *Expert Oracle Exadata*. 2015, Springer. p. 303-339.
23. Chow, W.S. and W. On Ha, *Determinants of the critical success factor of disaster recovery planning for information systems*. Information Management & Computer Security, 2009. **17**(3): p. 248-275.
24. McCormack, G.D. and R.F. Talaga Jr, *Multiple channel adaptive data recovery system*. 2002, Google Patents.
25. Gabber, E., et al., *Redundant data storage and data recovery system*. 2004, Google Patents.
26. Martin, B.C., *Disaster Recovery Plan Strategies and Processes*. SANS Reading, 2002.
27. V. S. Karwande, D.S.S.L., Prof. R. A. Auti, *The Data Recovery File System for Hadoop Cluster -Review Paper*. (IJCSIT) International Journal of Computer

- Science and Information Technologies, 2015. **6**(1): p. 365-367.
28. corporation, E., *MCD Data Domain Boost for Oracle Recovery Manager (RMAN)*. 2015: EMC Corporation.
29. Sharma, K. and K.R. Singh, *Online data back-up and disaster recovery techniques in cloud computing: A review*. International Journal of Engineering and Innovative Technology (IJEIT), 2012. **2**(5): p. 249-254.
30. Chivers, H. and C. Hargreaves, *Forensic data recovery from the Windows Search Database*. digital investigation, 2011. **7**(3): p. 114-126.
31. Puri, R., et al., *Forward error correction (FEC) codes based multiple description coding for Internet video streaming and multicast*. Signal Processing: Image Communication, 2001. **16**(8): p. 745-762.
32. Jadhao, M.V.G. and P.D. Gawande, *Performance Analysis of Linear Block Code, Convolution code and Concatenated code to Study Their Comparative Effectiveness*. 2012.
33. Toghuj, W. and G.I. Alkhatib, *Improved Algorithm for Error Correction*. International Journal of Information Technology and Web Engineering (IJITWE), 2011. **6**(1): p. 1-12.
34. Mitchell, G., *Investigation of Hamming, Reed-Solomon, and Turbo Forward Error Correcting Codes*. 2009, DTIC Document.
35. Jadhao, V.G. and P.D. Gawande, *Development Of Error Control Coding Technique In Digital Communication System*. International Journal of Managment, IT and Engineering, 2012. **2**(4): p. 304-316.
36. Jadhao, M.V.G.G., Prafulla D, *Performance Analysis of Linear Block Code, Convolution code and Concatenated code to Study Their Comparative Effectiveness*. 2012.
37. Moon Todd, K., *Error correction coding: mathematical methods and algorithms*. 2005 by John Wiley & Sons. 2005, ISBN 0-471-64800-0.
38. Yaakobi, E., et al. *Characterization and error-correcting codes for TLC flash memories*. in *Computing, Networking and Communications (ICNC), 2012 International Conference on*. 2012. IEEE.
39. Ryan, W. and S. Lin, *Channel codes: classical and modern*. 2009: Cambridge University Press.
40. Kumar, S. and R. Gupta, *Performance Comparison of Different Forward Error Correction Coding Techniques for Wireless Communication Systems*. 2011.
41. Morelos-Zaragoza, R.H., *The art of error correcting coding*. 2006: John Wiley & Sons.
42. Mrutu, S.I., A. Sam, and N.H. Mvungi, *Assessment of Non Transmittable Codewords Enhancement to Viterbi Algorithm Decoding*. International Journal of Communication Science and Information Security (IJCSIS), 2014. **12**: p. 9.
43. Singh, S. and S. Sujana. *ASIC Implementation Of Reed Solomon Codec For Burst Error Detection And Correction*. in *International Journal of Engineering Research and Technology*. 2013. ESRSA Publications.
44. Shrivastava, P. and U.P. Singh, *Error Detection and Correction Using Reed Solomon Codes*. International Journal of Advanced Research in Computer Science and Software Engineering, 2013. **3**.
45. Moon, Y.S.J. and N. Kaplan, *Introduction to Reed-Solomon Codes*. 2011, August.
46. van de Pavert, L., *REED-SOLOMON ENCODING AND DECODING*. 2011.
47. Revathy, M. and R. Saravanan, *HDL IMPLEMENTATION OF ALGEBRAIC SOFT DECISION ALGORITHM FOR RS CODES*. 2013.
48. Varsamou, M., I. Zacharias, and T. Antonakopoulos. *Correlated noise estimation and error correction in parallel storage channels*. in *Signal Processing and Information Technology (ISSPIT), 2013 IEEE International Symposium on*. 2013. IEEE.
49. Plank, J.S. and L. Xu. *Optimizing Cauchy Reed-Solomon codes for fault-tolerant network storage applications*. in *Network Computing and Applications, 2006. NCA 2006. Fifth IEEE International Symposium on*. 2006. IEEE.
50. Plank, J.S., *Optimizing Cauchy Reed-Solomon codes for fault-tolerant storage applications*. University of Tennessee, Tech. Rep. CS-05-569, 2005.
51. WIKIPEDIA. *Convolutional code*. 2015; Available from: https://en.wikipedia.org/wiki/Convolutional_code.
52. Mrutu, S.I., A. Sam, and N.H. Mvungi, *Forward Error Correction Convolutional Codes for RTAs' Networks: An Overview*. International Journal of Computer Network and Information Security (IJCNIS), 2014. **6**(7): p. 19.
53. NEHA, G.M.R., *CONVOLUTION ERROR CONTROL CODING -A REVIEW*, in *IRF International Conference*. 2015: Pune, India.
54. Garcia Planas, M.I., E.M. Souidi, and L.E. Um, *Decoding algorithm for convolutional codes under linear systems point of view*. 2014.
55. Grassl, M. and M. Rötteler. *Quantum block and convolutional codes from self-orthogonal product codes*. in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*. 2005. IEEE.
56. Raj, V., *Performance study on the suitability of Reed Solomon codes in communication systems*. CT International Journal of Information & Communication Technology, 2013. **1**(1).
57. Oenning, R.G.a.T., *Iterative Detection Read Channel Technology in Hard Disk Drives*. 2008: 2012 HGST, a Western Digital company.

58. Costello Jr, D.J., et al., *Applications of error-control coding*. Information Theory, IEEE Transactions on, 1998. **44**(6): p. 2531-2560.

Author Profile

Kilavo Hassan: Currently he is the Ph.D. student at Nelson Mandela African Institution of Science and Technology in the School of Computational and Communication Science and Engineering. He received his B.Sc. In Computer Science at the University of Dar Es Salaam, Tanzania and His M.Sc. in Information and Communication Science Engineering majoring Telecommunication Engineering at the Nelson Mandela African Institution of Science and Technology Arusha Tanzania. His research interest includes Forward Error Correction, Embedded System and Automation System.

Kisangiri Michael, received his Ph.D. degree in Telecommunication Engineering (2008) and Master of Science in Telecommunication Engineering (2002), both from Wroclaw University of Technology - Poland. Since then he has been working as academician with several universities. Currently he works with Nelson Mandela African Institution of Science and Technology located in Arusha, Tanzania. His research interests include evolutionary computation in Telecommunication Networks, Antenna design and Triangular mesh modeling.

Salehe I. Mrutu received his B.Sc. and M.Sc. degrees in Computer Science from The International University of Africa in 2003 and the University of Gezira in 2006 respectively. He received his Ph.D. in Information and Communication Science and Engineering in 2014 at Nelson Mandela African Institution of Science and Technology (NM-AIST) in Arusha, at the school of Computational and Communication Sciences and Engineering. He is currently serving as lecturer at the University of Dodoma (UDOM) under the school of Informatics since 2007. His research interests include forward error correction codes, algorithms analysis and designing, quality-of-service provisioning and resource management for multimedia communications networks.

Certificate based hybrid authentication for Bring Your Own Device (BYOD) in Wi-Fi enabled Environment

Upasana Raj

Student, Information Security and Cyber Forensics,
Department of Information Technology,
SRM University
Chennai, India

Monica Catherine S

Lecturer, Information Security and Cyber Forensics,
Department of Information Technology,
SRM University
Chennai, India

Abstract— Approval of the strategy, ‘Consumerization of IT’ by the organizations, does not only save money and increase business agility, but also improves employee productivity and satisfaction, lowers IT procurement, support costs and improves collaboration. Organizations have started to develop “Bring Your Own Device” (BYOD) policies to allow their employees to use their owned devices in the workplace. It’s a hard trend that will not only continue but will accelerate in the coming years. In this paper we focus on the potential attacks that can strike on BYOD when authenticating to a Wi-Fi network. It also enumerates the authentication protocols and methods. A proposal for stringent, indigenous hybrid authentication for a device that embraces BYOD strategy in a Wi-Fi enabled environment is proposed to the end of the paper.

Keywords— Bring Your Own Device (BYOD), Wi-Fi, Authentication, Authorization, Certificate.

I. INTRODUCTION

Recent years have noted the outburst in consumer mobile computing device, accompanied by their attractive user friendly and falling prices which makes these mobile devices requisite and within easy reach of the common man. Today, with even greater advances in consumer technology, mobile applications and the affordability of smart and powerful mobile devices, organizations are more challenged than ever to incorporate them into the enterprise IT architecture. The main concern of the IT department is that of security and data privacy risks that accompany with the BYOD movement with increased support costs.

BYOD not only holds the promise of not enabling companies to become more agile and customer focused, but also helping employees swiftly create and apply knowledge at work, which is key driving competitive lead in a knowledge-driven economy. The issue today is for the enterprise to embrace these changes in ways that improve organizational effectiveness and productivity while mitigating risks. The utmost concern of BYOD is the consequences of the usage of the unsecured personal mobile devices for handling corporate data.

‘Consumerization of IT’ is transmuting the traditional IT landscape of organizations and the way employees use technology for work purposes. Using their own devices help employees to handle the device in an efficient manner as they are more familiar and comfortable with functionalities of the device and also makes it possible for the employees to flexibly work from their home or on the road as per their convenience. Also using its own device means an employee will take an extra care for its safeguard. Since corporate information and personal information are on the same device, the ease of use to fetch information also enhances. Due to this the communication and work would be faster and efficient.

According to Gartner’s study published in Symantec report 2015 [14],

- 77% of the employees use their own phones for work.
- 74% enterprises allow employees to bring their own devices to work.
- The number of devices managed in the enterprise increasing by 72% from 2014-2015.

Security complements quality. Ignorance to the security aspect has disastrous effects on the organization leading to loss of confidential data and organization’s reputation. When employees try accessing Wi-Fi through the devices, a stringent authentication is required. Security concerns have held back Wi-Fi adoption in the corporate world. Hackers and security consultants have proved how vulnerable is the current security technology, wired equivalent privacy (WEP), used in most Wi-Fi connections. Data is the most crucial asset to any

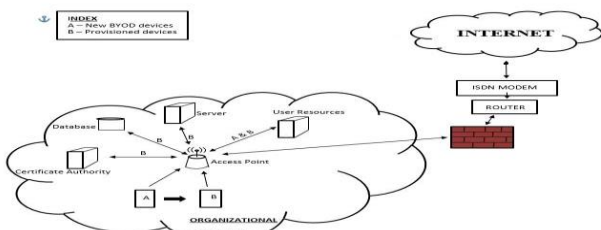


Figure No 01: BYOD Architecture

organization. With emergence of BYOD it has become more obligatory to implement a stringent, indigenous, hybrid authentication mechanism. The rest of the paper is organized as follows, Section II presents the literature survey for BYOD, introduction to attack that are pertinent to BYOD, authentication standards, methods and protocols practiced for the ‘consumerization of IT’. In section III we pen down the attacks, its mechanism, scenarios, symptoms and their consequences. In section IV a scheme is proposed for a stringent indigenous, hybrid authentication for BYOD strategy in a Wi-Fi enabled environment.

II. BRING YOUR OWN DEVICE (BYOD) TREND

A. Mobility Strategy in an organization

The four mobile strategies that come up are, here is your own device (HYOD), choose your own device (CYOD), Bring Your Own Device (BYOD), On Your Own Device (OYOD).

1) Here is your own device (HYOD)

The devices are provided by the organization. There is total control on the device by the enterprise. Organization provides total support for the device starting from installation to configuration and settings of the device [1].

2) Choose Your Own Device (CYOD)

The organization provides the employees with bunch of devices and employee makes his choice amongst them. Policies are not very stringent for these, they can install some specific applications and software [1].

3) Bring Your Own Device (BYOD)

Employee buys a device of his choice and pays for the same and uses it for his work at the organization. Policies are thus weaker and the organization has less control over these devices. Users can install applications as per their choice, if that complies with the policies and regulations of that organization [1].

4) On Your Own Device (OYOD)

The end user, i.e., the employee can buy and use any device. No policies are imposed.

BYOD can assure security as well as productivity if correct measures are taken, thus it is the recent trend seen in the corporate environment and is expected to flourish at a higher rate.

B. BYOD and its market overview

Bring your own device (BYOD) is a global phenomenon, 89 percent of IT departments enable BYOD in some form [15]. It offers advantages namely, enhanced productivity, increased revenue, reduced mobile costs, IT efficiencies.

With large numbers of employees already having smart technology, some organizations view this as an opportunity to implement new technology without having to pay for the devices themselves. Some major corporations and organizations choose to avoid changing their security protocols and migrate to BYOD because they do not want to risk the increased exposure to cyber

threats and data breaches. Another major reason why some corporations avoid switching to BYOD is because it is still relatively new and poses far numerous security threats, from a data security point of view, which could be found in the devices or even in their apps.

C. BYOD security risks[7]

- 1) Unified policy management
- 2) Securing and delivering corporate network access and services.
- 3) Device protection
- 4) Secure data transmission

D. Attacks pertinent to BYOD

Table No 01: Taxonomy of BYOD attacks [2]

| Component | Security Attacks | | |
|-----------|--|-------------------------|-------------------------------------|
| | Active | Passive | Privacy |
| User | - Man in Middle -Social Engineerin g | -Eavesdropping | Data privacy for company and client |
| Network | SSL Attack | | |
| Software | Malware APT | | |
| Physical | | Lost and stolen devices | |
| Web | SQL injection | | |

1) Social Engineering (SE)

Social engineering is the art of manipulating people so they give up confidential information.

2) Distributed Denial of Service (DDOS)

A denial of service is characterized by an explicit attempt by an attacker to prevent authenticate users from using computing of resources.

3) Inside Attack

An insider attack is a malicious attack perpetrated on a network or computer system by a person with authorized system access.

4) Man in middle attack (MITM)

A man-in-the-middle attack is an attack where the attacker secretly relays and possibly alters the communication between two parties who believe they are directly communicating with each other [2].

5) Secure Socket Layer attack (SSL)

The secure Socket layer attack is a type of attack that focus on breaching the vulnerabilities of the network protocol. SSL/TLS is the common protocol that targeted by the attacker for this attack [2].

III. AUTHENTICATION AND AUTHORIZATION

A. Authentication Protocol

Table No 02: Comparative analysis of authentication protocol [29]

| Sr No | Authentication | Features | Drawbacks |
|-------|----------------|---|---|
| 1 | PAP | Each user who wants to access the network should have a registered username and password to network access server. PAP is a point to point authentication protocol. | PAP is a weak authentication without any encryption. PAP is vulnerable to password guessing. |
| 2 | CHAP | CHAP is a three way handshake authentication protocol. Authentication is based on username and password. It can be used with other wireless authentication protocol such as TTLS. It has hash value method for security. | It is for point to point connection and it is not for such mesh network. CHAP originally designed for wired network. It is not as useful for large network since every possible is maintained at both ends of link. |
| 3 | Shared key | Shared key is a authentication protocol used in IEE 802.11. The shared key uses WEP or WPA for encryption, uses a four-way handshake authentication. | Shared one key for all users. No mutual authentication. Vulnerable to inside and packet spoofing attack. When changing the key one should advertise to all users. |
| 4 | EAP-802.1x | It is a port based authentication. It is a point to point authentication protocol. It is for single hope wireless network. It is flexible with different other protocols and can implement with new protocols. It is a data link layer protocol. It provides a way to dynamically send keys to clients. | It cannot support multi hop network. It is a framework not a protocol, this can found from its list of protocols. It needs a separate encryption protocol with self. |
| 5 | EAP-RADIUS | The communication of the RADIUS is on different layers of OSI protocol stack. It can use different authentication protocols such as EAP-802.1x, EAP-TTLS, EAP-MD5, PAP, CHAP, and etc... Four type of message exchange between authentications related devices. It uses UDP for data transmission. | The network that uses RADIUS has two parts, wired and wireless, and in the wired part data will be on clear text. It always needs another protocol and separate software for authentication, and encryption. |
| 6 | PANA | PANA developed by IETF network group. PANA carry the authentication between client and authentication server. PANA work on multi hop network and point to point network. It is very new and still under progress PANA works based on IP layer. PANA is designed for mutual authentication, and fast re-authentication. | The discovery and handshake phase is prone to spoofing attacks by a malicious node as there is no security relationship between PAA and PaC at that stage. Most of router firmware does not have this framework. |

B. Extended Authentication Protocol (EAP) Overview

EAP is not and does not specify any authentication methods, but by using frame work provided by IEEE 802.1x it supports various authentication methods. EAP supports multiple authentication methods, for example, one-time password, certificate, public key authentication, smart key, Kerberos. EAP authentication process can be explained as:

- 1) The authenticator, which sends request of authentication for supplicates
- 2) The user, responds ti each request.
- 3) After an innings of several Response/Request messages the authentication ends.

Authentication process is completed with a success or failure messages.

Table No 2: Comparison of authentication approaches under EAP [6]

| | EAP-MD5 | EAP-LEAP | EAP-FAST | EAP-TLS | EAP-PEAP | EAP-TTLS |
|----------------------------------|---|--------------------------------------|--|-------------------|--|--------------------|
| Implementation | Challenge Based | Password Based | PAC | Certificate Based | Server Certificate | Server Certificate |
| Authentication Attributes | Unilateral | Mutual | Mutual | Mutual | Mutual | Mutual |
| Deployment Difficulties | Easy | Easy | Easy to moderate depending on security | Hard | Moderate | Moderate |
| Dynamic Key Delivery | No | Yes | Yes | Yes | Yes | Yes |
| Server Certificate | No | No | Yes for maximum security | Yes | Yes | Yes |
| Supplicate Certificate | No | No | Yes | No | Yes | Yes |
| Tunneled | No | Yes | Yes | Yes | No | Yes |
| WAP Compatibility | Weak | Moderate | Weak to secure depending on implementation | Maximum Security | Secure | Secure |
| WLAN Security | Weak | Moderate | Weak to secure depending on implementation | Maximum Security | Secure | Secure |
| Vulnerabilities | Identity exposed, Dictionary attack, MITM attack. | Identity exposed, Dictionary attack. | Maximum security is comparable to PEAP and TTLS. | Identity exposed. | MITM attack, Identity hidden in phase 2 but potential exposure in phase 1. | MITM |

C. Authentication Methods

1) Two factor authentication

Two-factor authentication is a security process in which the user provides two means of identification from separate categories of credentials; one is typically a physical token, such as a card, and the other is typically something memorized, such as a security code.

2) Multi factor authentication (MFA)

Multifactor authentication is a security system that requires more than one method of authentication from independent categories of credentials to verify the user's identity for a login or other transaction.

3) Single Sign-on (SSO)

Single sign-on is a session/user authentication process that permits a user to enter one name and password in order to access multiple applications.

4) Public Key Infrastructure (PKI)

A public key infrastructure supports the distribution and identification of public encryption keys, enabling users and

computers to both securely exchange data over networks such as the Internet and verify the identity of the other party.

5) Digital Certificate

Digital certificate are basis for device authentication. Passwords are inherently vulnerable to phishing attacks, whereas user certificates are not [18].

a) General Certificate:

Digital Certificates are a means by which consumers and businesses can utilize the security applications of PKI. The certificate contains the name of the certificate holder, a serial number, expiration dates, a copy of the certificate holder's public key and the digital signature of the certificate-issuing authority (CA) so that a recipient can verify that the certificate is real.

b) Attribute Certificate:

Attribute Certificate is property set issued by Attribute Authority (AA) related to owner, it defines owner permissions in systems. AC is on basis of general certificate based on the public key infrastructure certificate, prevent the attacker to forge. AC format is defined by the X.509 version 3. This includes two main operations, i.e., to register a certificate for an

employee and to load the certificate when he requests for authentication, thus verifying and validating the device.

6) *One-time password*

An OTP is more secure than a static password, especially a user-created password, which is typically weak. OTPs may replace authentication login information or may be used in addition to it, to add another layer of security.

7) *Smart key*

Smart cards can provide personal identification, authentication, data storage, and application processing. Smart cards may provide strong security authentication for single sign-on (SSO) within large organizations.

8) *Kerberos*

Kerberos is a computer network authentication protocol which works on the basis of 'tickets' to allow nodes communicating over a non-secure network to prove their identity to one another in a secure manner. It aims primarily at a client-server model and it provides mutual authentication, both the user and the server verify each other's identity. Kerberos protocol messages are protected against eavesdropping and replay attacks. Disadvantages of Kerberos protocol,

- a) Single point of failure
- b) The administration protocol is not standardized and differs between server implementations
- c) If infrastructure of an organization requires an increase across a highly distributed networks, Kerberos doesn't have such scalability to be implemented.

9) *Secure Socket Layer*

Secure Sockets Layer protocol (SSL) is based on TCP/IP protocol, providing security for client and server communication. It uses asymmetric encryption technology to realize safety transfer of/from both sides information, this ensure the confidentiality and integrity of information, also can distinguish the identity of the conversation.

In SSL protocol handshake process, there is a process asking for each other's certificate for identity authentication which improves the safety of the connection but there still exists some drawbacks,

- 1) It cannot provide access control functions
- 2) Different users connect to the same server as they use the same authorization, which is a bad practical implementation practice.
- 3) It can only provide one-to-one SSL connection
- 4) Multi-level and multi-certificate chain of trust relationship cannot be achieved.

D. *Mobile Device Management (MDM)*

Mobile Device Management is a tool that in a centralized manner controls the devices and can do over the air configuration remotely to those devices that are connected to the network [5]. The MDM authenticates devices by exchanging certificates from the organization's server, defining access rights. MDM continuously sync's and stores backups of the data. All communication is generally secured by SSL/TLS to provide an encrypted channel.

E. *Device Fingerprinting*

Device fingerprinting allows devices to be identified, or fingerprinted, as an additional means of authentication.

F. *Device Encryption*

Encryption is a cornerstone of BYOD security. Encrypted VPNs, using IPSec provide for the confidentiality and integrity of data in transit, but this may leave data on the devices unencrypted.

IV. PROPOSED APPROACH

To avoid limitations in SSL protocol we propose a hybrid authentication system, combining use of attribute certificate with the general certificate to authenticate a device and the service and multi factor authentication to verify the user. When client requests for authentication it will be validated on basis of a general certificate and an attribute certificate. This allows to achieve access control along with multi-level and multiple certificate chain of trust relationship.

In order to mitigate DDOS attacks we propose the concept of three-tier captcha with username, password. Three tier captcha is implemented to mitigate and resist attacks like, dictionary attack, pixel count attack, pre-processing and vertical segmentation [5].

Advantages of a three tier captcha are,

- Enhances security
- Is easy to use
- Prevent automated attacks
- Difficult to identify the pattern

The basic flow of the authentication system will be as follows,

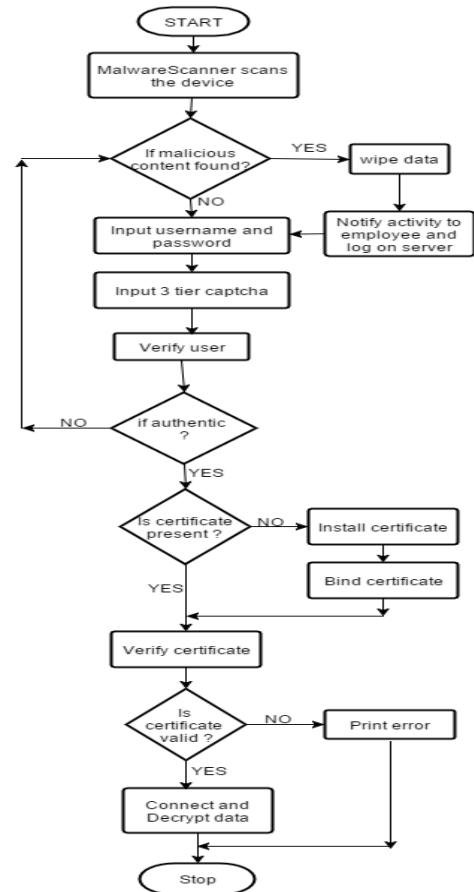


Fig 02: Flowchart for the authentication mechanism

Authentication mechanism uses Oracle 10g as its database for storing the employee's credentials. The User's personal information is stored in UserInfo table of database. When the user tries to connect to the AP of the organization's Wi-Fi, It runs a malware scan over the device to check signs of malicious activity by the MalwareScanner.

1) *MalwareScanner:*

Malware scanner checks for any malicious applications or services in the device before it connects to the organizations Wi-Fi.

2) *VerifyUser:*

When the user logs in, the server firstly verifies his username/password.

3) *InputThreeTierCaptcha:*

User needs to enter captcha based on the query displayed on screen. Server will verify the input, this will help mitigate DDOS attacks. This will help mitigate the DDOS attacks.

Algorithm of three-tier captcha

- 1) Create random string, combination of alphabets, numbers and symbols
- 2) Create an image of it with few noise
- 3) Generate a random query related to the code
- 4) Put the image and query onto the user interface
- 5) Allow user to provide the input
- 6) Examine input by user with value in the session
- 7) If input is correct, allow user to proceed
- 8) If input is incorrect, generate another captcha image and give limited chances, say 3, if in three attempts the user does not provide the correct username, password and captcha answer a text message is send to employees number/email account and an entry is made to server logs.

After verifying the user's username/password, authentication system will query user's information of certificate in the back-end database. If the user has applied for a digital certificate, the system will automatically jump to the certificate validation page.

Otherwise the system will automatically jump to the page of requesting certificates, and guide the user for installation of digital certificates.

4) *InstallCertificates:*

Installing certificates enables device fingerprinting, distribution of the Wi-Fi settings to the device used by the employee which helps him to access the Wi-Fi and the resources through it.

5) *BindCertificates:*

When the user applies for digital certificates successfully, the system will bind new digital certificates applied with the user's personal information. The user's personal information has bound with the digital certificate applied in the database. The workflow of binding the user's personal information with digital certificates is shown as follows:

- a) Download BindKey control to the local Web browser.
- b) Call function GetPK of the control to extract the public key of digital certificates, and send the public key to the server.
- c) After receiving the public key, the server enters the information of public key to the corresponding records in the

table. UserInfo according the username of the user logged in. Then the binding of personal information and the digital certificates completes.

6) *VerifyCertificate:*

After binding personal information with the digital certificates successfully or when the user logs in again, the authentication system will verify user's certificate, also sends key attribute values in attribute certificate to AA, AA verifies authenticity of attribute certificate, and determines client's role authority, returns the result back to server finally. When handshake process is over,

- a) User can communicate with server with role which attribute certificate aproves, server can ensure the security of the system through giving corresponding authority according to users' role.
- b) Server should ask user for certificate validation and attribute certificate validation at the same time.
- c) After user receives the request, it sends its public key, certificate and attribute certificate to server.
- d) Server verifies the certificate after receives it, then sends key attribute values in attribute certificate to AA, AA verifies authenticity of attribute certificate, and determines user's role authority, returns the result back to server finally.
- e) After the user and the server have verified each other, then to the following steps.
- f) When handshake process is over, user can communicate with server with role which attribute certificate gives it, server can ensure the security of the system through giving corresponding authority according to users' role.

The attribute certificate added to SSL protocol handshake process describes as figure below:

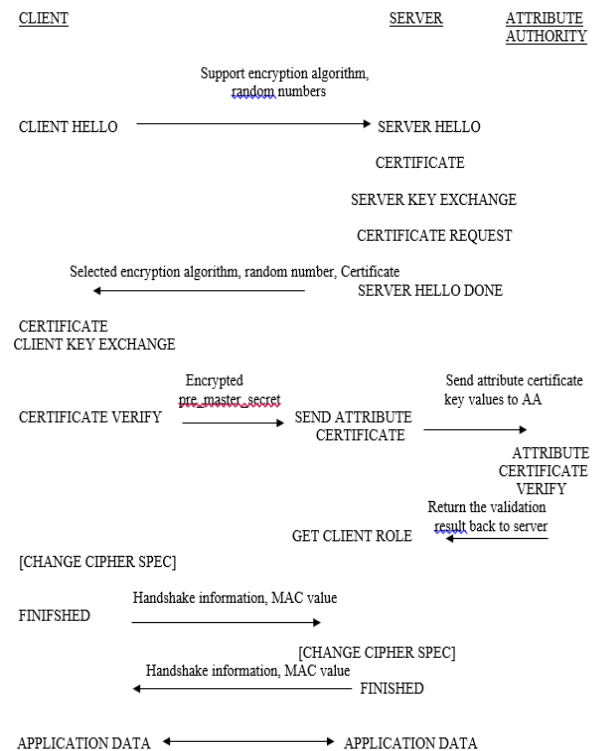


Fig 02: Attribute certificate mechanism [4]

7) *DataDecrypter*:

Once the device, user is authenticated and authorized for the given service the confidential data gets decrypted.

8) *DataEncrypter*:

When the employee, disconnects the organizational Wi-Fi, the data termed under confidential and critical are encrypted.

CONCLUSION

A hybrid authentication system is required to ensure productivity and efficiency adopting BYOD trend. A malware scanner scans the device when the device tries to connect to the organization's Wi-Fi. A three tier captcha collective with username and password, will authenticate and verify the user. Attribute certificate mechanism in SSL enables multi-level trust relationship. Attribute certificate has certificate signer unique identifier and extensions, can add certificate signer's certificate in extensions, when users' own certificate cannot be verified by the other side, they can call extensions of attribute certificate to ask for certificate of certificate signer, which can achieve the goal of multi-level authentication.

ACKNOWLEDGMENT

The research presented is being performed for the SRM University, Chennai, India in partial fulfillment of the requirements for the Master of Technology degree. I take this opportunity to express my profound gratitude and deep regards to Dr. S. Magesh for his exemplary guidance, monitoring and constant encouragement. Prof S. Rajendran, for their patience and 'never-give-up' attitude with their students. I am thankful to all lecturers at SRM University for their support and comments. I am obliged to all my professors, teachers at SCOE, Pune, for the faith, encouragement and valuable guidance.

REFERENCES

[1] Prashant Kumar Gajar, Arnab Ghosh, Shashikant Rai, "Bring your own device (BYOD), Security risks and mitigating strategies", Journal of Global research in Computer Science, Volume 4, April 2014.
[2] Manmeet Mahinderjit Singh, Soh Sin Siang, Oh Ying San, "Security attacks taxonomy on Bring your own model (BYOD)", International Journal of Mobile Communication and Telematics, (IJMNCT) Vol. 4, No.5, October 2014.
[3] Khoula AlHarthy, Wael Shawkat "Implement network security control solutions in BYOD environment", IEEE International Conference on Control System, Computing and Engineering, 2013.
[4] LI Wei, XIANG Shuyue, CHEN Shuangbao, "Improvement method of SSL protocol Identity authentication based on attribute certificate", IEEE, 2012.
[5] Bharat Yadav, et.al, "Introducing three-tier captcha to prevent DDOS attack in Cloud computing", International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169, Volume: 2 Issue: 7, 2014.
[6] "Authentication methods for Wi-Fi networks", International Journal of Application or Innovation in Engineering and Management (IJAIEM), volume 1, Issue 3, March 2013.

[7] Robert J. Maverick, et.al, "Legal issues with corporate BYOD programs", SANS Institute InfoSec Reading Room, 15th may, 2012.
[8] Aaron M French, et.al, "Current status, issues and future of Bring your own device (BYOD)", Communications of the Association For Information Security, Article 10, Volume 35, pp. 191-197, November 2014.
[9] Jeff Loucks, Richard Medcalf, Lauren Buckalew, Fabio Faria, "Symantec BYOD report", CISCO IBSG Horizons, 2015.
[10] Kiyohide NAKAUCHI, et.al, "Bring Your Own Network – Design and Implementation of a Virtualized Wi-Fi Network", 11th Annual IEEE CCNC, Smart Spaces and Wireless Networks, 2014.
[11] Nima Zahadat, Paul Blessner, Timothy Blackburn, Bill A. Olson, "BYOD security engineering, framework and its analysis", Computers & Security, 2015.
[12] LI Fen, LIU Quan, PANG Liaojun, PEI Qingqi, "Identity Authentication based on fuzzy vault and digital certification", International Conference on Computational Intelligence and Security, 2010.
[13] Wu Kehe, Chen Wei, Ge Yueguang, "The research and implementation of the authentication technology based on digital certificates", Fourth International Conference on Computational and Information Sciences, 2012.
[14] Richard Medcalf, Lauren Buckalew, Fabio Faria "Meeting mobile and BYOD security challenges with digital certificates", Symantec, 2015.
[15] Joseph Bradley, Jeff Loucks, James Macauley, et.al, "Bring your own device (BYOD): A Global Perspective", Survey report by Cisco Horizon, April 2013.
[17] Hao Liu, Hideaki Goto, "Certificate-based, Disruption-tolerant Authentication System with Automatic CA Certificate Distribution for eduroam 38th Annual International Computers, Software and Applications Conference Workshops, IEEE, 2014.
[18] Mao Xiaodong, Wang Dahu, Zhong Tong, Huang Keming "Research on a certificate-based authentication for mobile ad-hoc networks", IEEE International Symposium on Computer, Communication, Control and Automation, 2010.
[19] Sebastian Koffer, Kevin Ortbach, Iris Junglas, et.al, "The influence of IT consumerization on individual IT innovation behaviour", Business and Information System Engineering, volume 57, Issue 6, Springer, 2015.
[20] Deepak Sanghara, Vishal Gupta, "Exploring security approach in BYOD environment", Advanced Computing, Networking and Informatics- Volume 2, Volume 28, Springer, 2015.
[21] Wan Ng, "Mobile Learning: Bring Your Own Device and personalized device", New Digital Technology in Education, Springer, 2014.
[22] D. Jaramillo, N. Katz, B. Bodin, et.al, "Cooperative solutions for bring your own device (BYOD)", IBM journal of research and development, volume 57, Issue 6, ACM, 2013.
[23] Aessandro Armando, Gabriele Costa et.al, "Bring your own device, securely", Applied Computing, 2013.
[24] George F. Hurlburt, Change Index, "BYOD: Security and Privacy considerations.", IEEE Computer Society, Issue No 5, 2012.
[25] Scarfo A, "New security perspective around BYOD", Broadband, Wireless Computing, Communication and Applications (BWCCA), IEEE, 2012.
[26] Lennon R G, "Changing user attitudes to security in bring your own device and cloud", Cloud & High Performance Computing Science, IEEE, 2012.
[27] Eslahi M, Naseri M V, Hashim H, "BYOD : current state and security challenges", Computer Applications and Industrial Electronics (ISCAIE), IEEE, 2014.
[28] Neal Leavitt, "Today's mobile security requires a new approach", Technology news, IEEE Computer Society, IEEE, 2013.
[29] Jani Suomalainen, Jukka Valkonen, N. Asokan, "Security Associations in Personal Networks: A Comparative Analysis", Security and Privacy in Ad-hoc and Sensor Networks, Springer, 2007.

Impact of Spectrum Sensing and Primary User Activity on MAC Delay in Cognitive Radio Networks

Elham Shahab

Department of Information Technology Engineering
Nooretouba Virtual University of Tehran
Tehran, Iran

Abstract—Cognitive radio (CR) technology is used in wireless networks with the aim of handling the wireless spectrum scarcity. On the other hand, using CR in wireless networks has some new challenges and performance overheads that should be resolved. It is crucial to study the effects of the unique characteristics of cognitive radio over the various protocols. In this paper, a simulation-based performance evaluation is presented in the term of MAC delay in a CR user. The effects of spectrum sensing parameters (sensing duration and frequency), the primary users' activity and the number of spectrum channels are investigated on the delay of MAC layer. The growth and decay rates of MAC delay are studied more in detail through the various NS2-based simulations. The results give some fruitful insights for formulating the delay of MAC based on the CR unique parameters.

Keywords—Cognitive radio; MAC; Delay; Primary user; Spectrum sensing

I. INTRODUCTION

Taking the advantage of cognitive radio (CR) technology in order to use the wireless spectrum dynamically has been widely used in wireless networks [1]. Cognitive radio enables the dynamic spectrum access (DSA) approach in wireless networks [2]. Wireless networks with CR-enabled nodes are generally called cognitive radio networks (CRNs) [3]. A CR node operates based on a cognitive radio cycle [4]. The cognitive radio cycle composed of four main stages: spectrum sensing, spectrum decision, spectrum sharing and spectrum hand-off [5]. A CR node is allowed to communicate on a wireless channel in the absence of primary user (PU) who has priority to use the spectrum channels [6]. The unique characteristics of CRNs, i.e., periodic spectrum sensing and operating in the absence of primary user, make new challenges in the modeling and performance evaluation of various protocols in MAC, network and transport layers. The performance evaluation of different protocols in CRNs can be fruitful in order to investigate the overhead of using CR technology in wireless networks. Therefore, being aware of the CR overhead in wireless networks can lead to design some efficient protocols for CRNs.

There are a lot of research activities in the literature about the performance evaluation of various protocols in CRNs. Researchers have evaluated the performance of CRNs in the

terms of diverse factors such as throughput, packet loss, delay and jitter.

In [18], the authors model and evaluate the sending rate distribution of source nodes in the transport layer of CRNs. The [10] establishes a balance between packet delay and sensing time through proposing a MAC protocol. In [15], an admission control is proposed in order to support delay sensitive communications of CR users in cognitive radio networks. Authors of [9] propose an admission control with the aim of minimizing the end-to-end delay and jitter in CRNs. Expected packet delay in CRNs is modeled based on waiting and transmission delays in [17]. The authors of [21] propose two optimal scheduling methods in order to maximize the throughput and minimize the scheduling delay in CRNs. The stochastic delay and backlog bounds of transport layer in CRNs are modeled in [16]. In [13], the optimality of congestion control schemes in cognitive radio networks are investigated in order to minimize the congestion probability and the delay of transport layer. The authors of [20] investigate the challenges of delay-sensitive data transport based on the unique features of cognitive radio. In [14], [11] and [19], the effect of dynamic spectrum access on the throughput of transport layer is studied in cognitive radio networks. The impact of CRNs characteristics over the performance of routing and transport protocols is evaluated based on simulations in [12]. The packet loss probability, end-to-end delay and throughput are evaluated based on the cognitive radio parameters in [22].

Although there are several studies over the performance evaluation of CRNs, it is needed to study on the delay behavior of MAC layer for a CR node individually. Investigating the behavior of MAC delay for a CR node can give us an appropriate insight to

- design the efficient MAC protocols for CRNs and
- adjust the existing MAC protocols and CR-related parameters

with the aim of improving the MAC delay performance of CRNs.

In this paper, we propose a simulation-based study on the MAC delay behavior of a CR node. The impact of the spectrum sensing, primary user activity and the number of wireless channels on the MAC delay is investigated by various simulations. The increasing/decreasing rate of MAC delay is

studied and compared based on the changing of various parameters. The results of this paper about the behavior of MAC delay can be helpful in order to formulate the delay of MAC with regard to CR unique characteristics. Also, it can be used to have some fruitful insights to determine the appropriate amounts of sensing duration, sensing frequency, the number of wireless channels in CRNs with the aim of minimizing the MAC delay.

In the rest of paper, the Section II describes the system model. The Section III studies the performance of MAC delay based on the sensing duration and frequency (Section III-A), the number of wireless channels (Section III-C) and the PU activity (Section III-B). Finally, the Section IV concludes the paper. Ease of Use

II. SYSTEM MODEL

A. Spectrum Sensing

A CR node needs to sense the wireless spectrum for a predefined duration (sensing time) and check the presence of primary user (PU) who has the license to use the spectrum channels. If there is no PU on a wireless channel, the CR node is allowed to communicate on the free wireless channel for a specific time (operating time). Since a CR node cannot sense the spectrum and send data on it simultaneously, the spectrum sensing is done periodically with a specific period with the aim of minimizing the amount of interference between the communication of CR nodes and the PU [7]. The ideal sensing is assumed without any errors in detection PU presence.

Let t_s and f_s be the spectrum sensing duration and frequency, respectively. In the other words, a CR node senses a spectrum channel for t_s with the period of $\frac{1}{f_s}$. Each period is composed of two durations. The first one is named the spectrum sensing duration and the second one is called the operating duration. Let t_o be the operating duration of a CR node at each period that is equal to $t_o = \frac{1}{f_s} - t_s$.

B. Primary User Activity

Primary users have higher priority to use the wireless spectrum channels. Modeling of PU activity has a high degree of importance because of its impact on the communications of CR users. The most common model for PU activity is the two-state birth/death Markovian process [8] with birth rate of β and death rate of α [7]. In CRNs, two states of Markovian process are named ON and OFF states. In the ON state, the wireless channel is busy by the primary user (the PU is ON (active) on the channel). In the OFF state, the PU is not active on the channel. The birth rate (β) and death rate (α) are called the entrance and departure rates of primary users, respectively.

C. Wireless Channels

Let N be the number of wireless channels. For each channel, there is a licensed primary user that enters the channel with the mean entrance rate of β and leaves it with the mean departure rate of α .

III. PERFORMANCE EVALUATION OF MAC DELAY

In this section, the delay of MAC is investigated based on the various parameters, i.e., spectrum sensing duration and period, the number of wireless channels and the entrance and the departure rates of primary users. Our simulations are done based on NS2 [23] and CogNS [22] that is an NS2-based simulation framework for cognitive radio networks. The default simulation setup is illustrated in Table I. The default number of wireless channels is 1. The sensing and operating durations are considered 0.01 sec and 0.6 sec, respectively. The activity parameters of primary users are determined $(\alpha, \beta) = (3, 1)$. The packet size is selected 512 bytes. The duration of simulations are considered 500 sec.

TABLE I. DEFAULT SIMULATION SETTINGS

| Parameter | Default Value |
|-------------------------------------|---------------|
| The number of wireless channels (N) | 1 |
| Spectrum sensing duration (t_s) | 0.01 sec |
| Operating duration (t_o) | 0.6 sec |
| PU departure rate (α) | 3 |
| PU entrance rate (β) | 1 |
| Packet size | 512 bytes |
| Simulation time | 500 sec |

A. Spectrum sensing duration and frequency

Spectrum sensing duration and frequency are two basic factors of cognitive radio that have significant influence on the performance of CRNs and the quality of service (QoS) of primary users. Based on the amounts of sensing duration (t_s) and operating duration (t_o), sensing efficiency can be defined as follows [7]:

$$\varepsilon = \frac{t_o}{t_s + t_o} \quad (1)$$

Maximizing the amount of sensing efficiency increases the spectrum utilization of CR users. Therefore, the maximizing of sensing efficiency is an objective in CRNs in order to reduce the amount of interference on primary users. The amount of ε tends to 1 when either the amount of t_s tends to zero or the amount of t_o tends to ∞ . On the other hand, the small values of t_s leads to low PU detection probability that increases the amount of interference on primary users. Also, the decreasing of t_s increases the false alarm probability that reduces the spectrum utilization of CR users. From the point of view of CR users, increasing the value of t_o raises the sensing efficiency. However, high value of t_o can increase the interference on PUs. Therefore, the selection of optimal sensing and operating durations is a crucial task in CRNs [7].

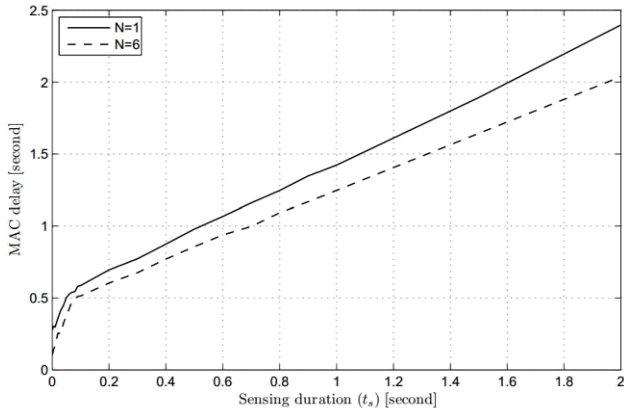


Figure 1. The behavior of MAC delay based on the changes of spectrum sensing duration

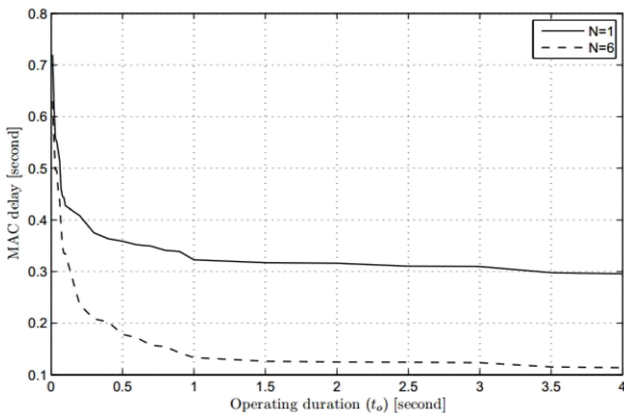


Figure 2. The behavior of MAC delay based on the changes of operating duration

Studying the impact of sensing and operating over the various factors in CRNs can help us to find the appropriate values of t_s and t_o . In this paper, we focus on the delay of MAC layer in a CR node. Generally, we expect that the decreasing of t_s and the increasing of t_o reduce the MAC delay of a CR node. On the other hand, the small amounts of t_s and the large amounts of t_o can violate the main objectives of cognitive radio networks.

We investigate the behavior of MAC delay per the different values of t_s and the various number of wireless channels. The value of t_s varies from 0.01 second to 2 seconds. The MAC delay is calculated based on the values of t_s for the number of channels (N) of 1 and 6. All of other parameters are considered as mentioned in Table I. The results are depicted in Fig. 1. As seen in Fig. 1, increasing the value of t_s raises the value of MAC delay. In this figure, the solid and dashed curves represent the MAC delay values for N=1 and 6, respectively.

According to the Fig. 1, the delay of MAC increases linearly with t_s independent of the number of channels. The slope of curves is changed around the $t_s=0.08$ second for both of N=1 and 6. The growth rate of MAC delay (the slope of curve) between the $t_s=0.01$ sec and $t_s=0.08$ sec is more than

the growth rate between the $t_s=0.08$ sec and $t_s=2.0$ sec independent of the number of wireless channels. As seen in Fig. 1, the growth rate of MAC delay with t_s is different for N=1 and 6. For N=6, the slope of curve is smaller. As a consequence, the changing of the sensing duration has more impact on MAC delay when there are the smaller number of wireless channels.

The behavior of MAC delay versus the values of t_o from 0.01 to 4 seconds is depicted in Fig. 2. The MAC delay is calculated for the number of channels (N) of 1 and 6. All of other parameters are considered based on the values of Table I. Based on the Fig. 2, increasing the value of t_o reduces the value of MAC delay. The solid and dashed curves represent the MAC delay values for N=1 and 6, respectively.

Based on the Fig. 2, a non-linear decrease of MAC delay is observed with to for different numbers of wireless channels. The decade rate of MAC delay with t_o is different for the various number of channels so that the decade rate for N=6 is greater than the decade rate for N=1. As a consequence, the changing of the operating duration has more impact on MAC delay when there are the greater number of wireless channels.

B. Entrance and departure rate of PU

The activity of primary users on wireless channels has a great impact on the performance of CR users. Based on the activity model of primary users, the probability of the presence a PU on a channel is as follows [7]:

$$P_{ON} = \frac{\beta}{\alpha + \beta} \quad (2)$$

where the α and β are the mean departure and entrance rates of primary users, respectively. The probability that a CR user can find a channel free of primary users can be calculated as follows:

$$P_{free} = 1 - (P_{ON})^N \quad (3)$$

where N is the number of wireless channels. Generally, we expect that the large amounts of β and the small values of α increase the MAC delay of a CR user.

In order to investigate the effect of α on MAC delay more in detail, we study the behavior of MAC delay versus the different values of α and for the various values of β . We change the value of α from 1 to 6 and study the behavior of MAC delay. This simulation is repeated for three different values of $\beta=1, 2$ and 3. All of other parameters are determined as explained in Table I. The results are depicted in Fig. 3. As it can be seen in the Fig. 3, the increasing of α reduces the MAC delay.

According to the Fig. 3, a non-linear decay of MAC delay can be seen with α for all values of β . The decay rate of MAC delay with α is different for the different values of β . For larger values of β , the rate of decay is larger. As a consequence, the changing of α has more effect on the MAC delay when the mean entrance rate of primary users is greater.

In order to study the impact of β on MAC delay more in detail, we consider the behavior of MAC delay versus the different values of β and for the various values of α . We change the value of β from 1 to 5 and investigate the behavior

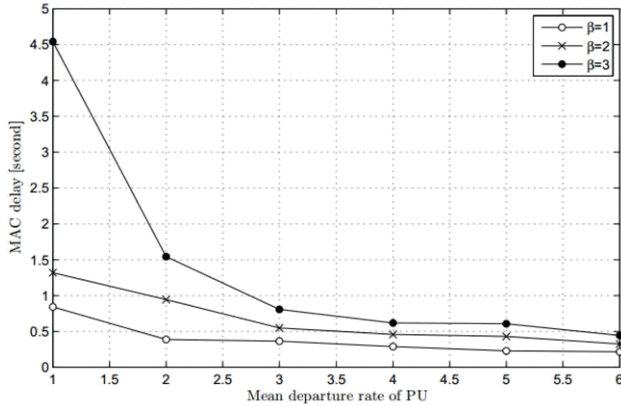


Figure 3. The behavior of MAC delay based on the changes of mean departure rate of PU

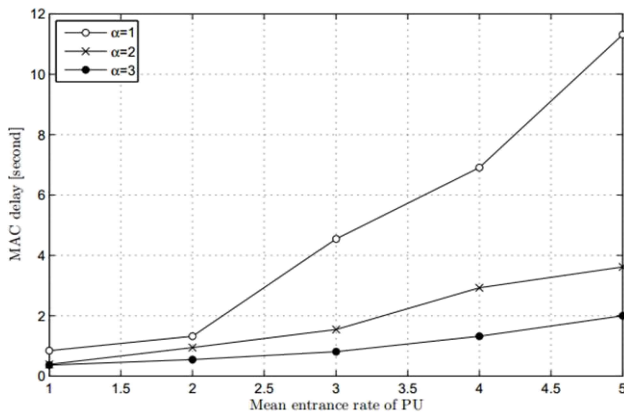


Figure 4. The behavior of MAC delay based on the changes of mean entrance rate of PU

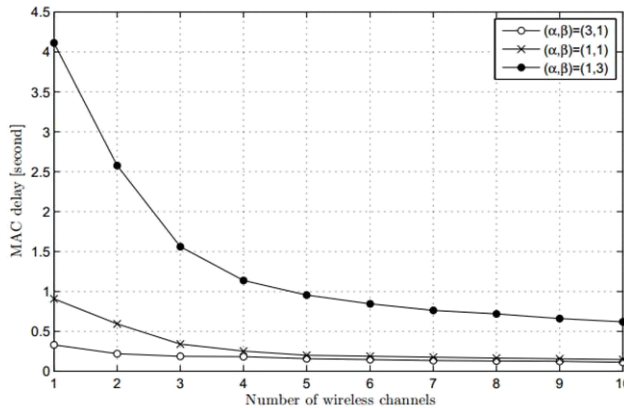


Figure 5. The behavior of MAC delay based on the changes of number of wireless channels

of MAC delay. This simulation is repeated for three different values of $\alpha = 1, 2$ and 3 . All of other parameters are determined as explained in Table I. The results are depicted in Fig. 4. As it can be seen in the Fig. 4, the increasing of β raises the MAC delay.

Based on the Fig. 4, a non-linear growth of MAC delay can be seen with β for all values of α . The growth rate of MAC delay with β is different for the various values of α . For smaller values of α , the growth rate is larger. As a result, the changing of β has more impact on the MAC delay when the mean departure rate of PU is smaller.

C. The Number of Wireless Channels

The number of wireless channels has an inevitable effect on the MAC delay of CR users. The larger number of wireless channels increases the chance of a CR user to find a channel free of primary users because according to the (3), the larger value of N decreases the amount of the term $(P_{ON})^N$ as a consequence the value of P_{free} increases.

In order to study more in detail the impact of the number of wireless channels on the MAC delay of a CR user, we evaluate the amount of MAC delay versus various values of N varying from 1 to 11 for the diverse activities of primary users, i.e., $(\alpha, \beta)=(1,3)$, $(\alpha, \beta)=(1,1)$, and $(\alpha, \beta)=(3,1)$. The related curves are illustrated in Fig. 5. As seen in this figure, a nonlinear decrease in the amount of MAC delay is observed with the increasing the value of N for all considered activities of primary users. The decay rate is larger for more active primary users.

IV. CONCLUSION

The behavior of MAC layer of a CR node with regard to the sensing duration, the operating duration, the activity of primary users and the number of wireless channels investigated through various simulations. The MAC delay increases linearly with increasing the value of sensing duration. By increasing the operating duration, the delay of MAC decreases non-linearly. The reducing of the mean entrance rate of primary users leads to a sharp decline (non-linearly) in the MAC delay. The increasing of the mean departure rate of primary users reduces the MAC delay. However, the decade rate of MAC delay due to the reducing of the PUs' entrance rate is more than the increasing departure rate of primary users. Moreover, the MAC delay goes up non-linearly with lessening the number of wireless channels.

REFERENCES

- [1] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks*, vol. 50, no. 13, pp. 2127-2159, Sep. 2006.
- [2] J. Marinho and E. Monteiro, "Cognitive radio: survey on communication protocols, spectrum decision issues, and future research directions," *Wireless Networks*, vol. 18, no. 2, pp. 147-164, Oct. 2011.
- [3] I. F. Akyildiz, W.-Y. Lee, and K. R. Chowdhury, "CRAHNs: Cognitive Radio Ad Hoc Networks," *Ad Hoc Networks*, vol. 7, no. 5, pp. 810-836, Jul. 2009.
- [4] J. Mitola, J. Maguire, G.Q., "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, col. 6, no. 4, pp. 13-18, 1999.
- [5] I. F. Akyildiz, W.-Y. Lee, and K. R. Chowdhury, "Spectrum management in cognitive radio ad hoc networks," *IEEE Network*, vol. 23, no. 4, pp. 6-12, Jul. 2009.
- [6] H. S. Kim, W. Ejaz, K. Al-Begain, A.-S. K. Pathan, N. U. Hasan, "Advances in cognitive radio sensor networks," *International Journal of Distributed Sensor Networks*, Article ID 631624, 2014.

- [7] W.-Y. Lee, I. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 10, pp. 3845-3857, 2008.
- [8] K. S. Trivedi, *Probability and Statistics with Reliability, Queuing, and Computer Science Applications*, 2nd Edition, Wiley-Interscience, 2001.
- [9] E. S. Hosseini, V. Esmaealzadeh, R. Berangi, and O. B. Akan, "A correlation-based and spectrum-aware admission control mechanism for multimedia streaming in cognitive radio sensor networks," *International Journal of Communication Systems*, in press, May 2015.
- [10] C.-L. Wang, H.-W. Chen, and Y.-X. Cheng, "Sensing-Delay Tradeoff for Cognitive Radio Networks with QoS Considerations," in *78th IEEE Vehicular Technology Conference (VTC Fall)*, 2013, pp. 1-5.
- [11] A. M. R. Slingerland, P. Pawelczak, R. Venkatesha Prasad, A. Lo, and R. Hekmat, "Performance of Transport Control Protocol Over Dynamic Spectrum Access Links," in *2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, 2007, pp. 486-495.
- [12] M. Di Felice, K. R. Chowdhury, W. Kim, A. Kassler, and L. Bononi, "End-to-end protocols for Cognitive Radio Ad Hoc Networks: An evaluation study," *Performance Evaluation*, vol. 68, no. 9, pp. 859-875, Sep. 2011.
- [13] V. Esmaealzadeh and R. Berangi, "On the Optimality of Generic Rate-based AIMD and AIAD Congestion Control Schemes in Cognitive Radio Sensor Networks," *International Journal of Distributed Sensor Networks*, vol. 2015, Article ID 614643, 2015.
- [14] Y. R. Kondareddy and P. Agrawal, "Effect of Dynamic Spectrum Access on Transport Control Protocol Performance," in *IEEE Global Telecommunications Conference (GLOBECOM)*, 2009, pp. 1-6.
- [15] F. Wang, J. Huang, and Y. Zhao, "Delay Sensitive Communications over Cognitive Radio Networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 4, pp. 1402-1411, Apr. 2012.
- [16] V. Esmaealzadeh, R. Berangi, E. S. Hosseini, and O. B. Akan, "Stochastic Backlog and Delay Bounds of Generic Rate-based AIMD Congestion Control Scheme in Cognitive Radio Sensor Networks," *Pervasive and Mobile Computing*, vol. 22, no. C, pp. 46-57, Sep. 2015.
- [17] W. Li, T. Song, X. Cheng, T. Jing, X. Xing, and T. Znati, "Delay modeling in mobile cognitive radio networks," in *2012 1st IEEE International Conference on Communications in China (ICCC)*, 2012, pp. 30-35.
- [18] V. Esmaealzadeh, E. S. Hosseini, R. Berangi, and O. B. Akan, "Modeling of Rate-based Congestion Control Schemes in Cognitive Radio Sensor Networks," *Ad Hoc Networks*, vol. 36, Part 1, pp. 177-188, Jan. 2016.
- [19] T. Issariyakul, L. S. Pillutla, and V. Krishnamurthy, "Tuning radio resource in an overlay cognitive radio network for TCP: Greed isn't good," *IEEE Communications Magazine*, vol. 47, no. 7, pp. 57-63, Jul. 2009.
- [20] A. O. Bicen, V. C. Gungor, and O. B. Akan, "Delay-sensitive and multimedia communication in cognitive radio sensor networks," *Ad Hoc Networks*, vol. 10, no. 5, pp. 816-830, Jul. 2012.
- [21] D. Gozuek and F. Alagoz, "Throughput and delay optimal scheduling in cognitive radio networks under interference temperature constraints," *Journal of Communications and Networks*, vol. 11, no. 2, pp. 148-156, Apr. 2009.
- [22] V. Esmaealzadeh, R. Berangi, S. M. Sebt, E. S. Hosseini, and M. Parsinia, "CogNS: A Simulation Framework for Cognitive Radio Networks," *Wireless Personal Communications*, vol. 72, no. 4, pp. 2849-2865, Apr. 2013.
- [23] Network simulator version 2. (<http://www.isi.edu/nsnam/ns/>).

A Survey of Receiver Designs for Cooperative Diversity in the Presence of Frequency Offset

Sylvia Ong Ai Ling, Hushairi Zen, Al-Khalid b Hj Othman, Mahmood Adnan

Abstract—Cooperative diversity is becoming a potential solution for future wireless communication networks due to its capability to form virtual antenna arrays for each node (i.e. user). In cooperative networks, the nodes are able to relay the information between the source and the desired destination. However, the performance of the networks (for instance – mobile networks, ad-hoc networks and vehicular networks) is generally affected by the mobility of the nodes. As the nodes' mobility rapidly increases, the networks are subjected to frequency offset and unknown channel properties of the communication links which degrades the system's performance. In a practical scenario, it is a challenging task and impractical for the relay and destination to estimate the frequency offset and channel coefficient especially in time-varying environment. In this manuscript, a comprehensive survey of existing literature for receiver designs based on Double Differential (DD) transmission and Multiple Symbol Detection (MSD) approach is presented to eliminate the complex channel and frequency offset estimation.

Index Terms—Cooperative Diversity, Double Differential, Frequency Offset, Multiple Symbol Differential Detection.

I. INTRODUCTION

IN recent years, wireless cooperative diversity has gained significant attention because of its capability to achieve low Bit Error Rate (BER), high network throughput, high data transmission reliability as well as spectral efficiency to support the demand for the rapid growth in wireless communications. By exploiting the broadcast nature and diversity gain of the wireless communication, cooperation between users can be realized without the requirement of physical antenna arrays being installed at the transmitting and receiving nodes [1,2]. The idea behind this technique is to enable the source to broadcast signals following independent wireless path towards its destination with the help of other node(s) that act as relay(s) in the transmission schemes such as One-Way Relay Network (OWRN) or Two-Way Relay Network (TWRN) in [3] as depicted Fig. 1. For OWRN, the transmission phase is divided into the broadcast and relayed phase. In the broadcast phase, the source broadcasts its information via a relay and directly towards the destination.

The source then remains idle during the second (i.e. relayed) phase. Simultaneously, the relay processes the received signals and relays the processed signals to its intended destination. In TWRN, when two nodes are communicating with each other through a relay, the relay processes the superimposed (summed) of the received signal from both nodes and broadcasts it back to its corresponding node [3,4].

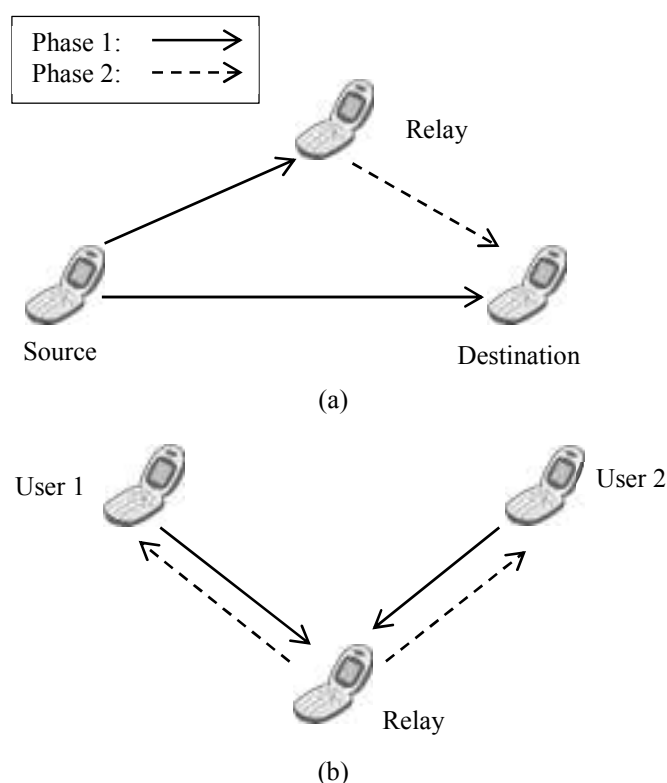


Figure 1: (a) One-Way Relay Network (OWRN) and (b) Two-Way Relay Network (TWRN) [3]

The received signal at the relay is processed based on several relaying protocols, such as Decode-and-Forward (DF) and Amplify-and-Forward (AF) [2,5]. For DF protocol, the relay decodes the source information and encodes the signals before retransmitting the information to the desired destination. However, the DF protocol may suffer from the error propagation problem constraint by the transmission of erroneous signal which deteriorates the whole system. On the contrary, in AF protocol, the relay receives the source information and simply amplifies the signals (both information and noise) with certain multiplication factors before retransmitting the scaled version of the signals towards the destination. The AF

*Sylvia Ong Ai Ling, Hushairi Zen, Al-Khalid b Hj Othman, Mahmood Adnan are with the Faculty of Engineering, Universiti Malaysia Sarawak, Kota Samarahan, Kuching, Sarawak, Malaysia. (e-mail: 15010028@siswa.unimas.my)

protocol is preferred over DF protocol due to its simplicity and importantly suitable for low-power consumption networks without error propagation [6].

At the destination, all the relayed signals and directly transmitted information is detected and combined using various diversity combining techniques in order to improve the performance of the faded received signals. Most of the existing literature assumes that the perfect channel coefficients knowledge is known at the relay and destination. But, in practical applications, it is a difficult and challenging task to perform perfect channel estimation due to the variation effects of the channel, especially in the fast fading environment. Hence, differential transmission scheme is studied in [7–9] by taking the advantages of eliminating channel knowledge in the networks or in situations wherein the channel information is unavailable. However, in a real-world wireless mobile network with the presence of frequency offset caused by the Doppler Shift – changes in frequency when a node moves toward source, the differential detection may fail to attain the desired performance. Moreover, due to the mobility of the nodes, the oscillators of the source, relay(s) and destination experience are experiencing difficulties to synchronize perfectly [10].

Generally, there are two major approaches to deal with the presence of frequency offset in wireless communication systems. One of the approaches to alleviate the frequency offset is to design tremendous estimators that estimate the frequency offset. The estimated offset is then compensated by way of implementing tracing circuit as well as frequency acquisition [11]. These, however, demonstrate a significant reduction of data rates owing to the transmission of pilot symbols as a reference in the estimation process, and increase the computational complexity. Thus, Double-Differential (DD) transmission is devised in [12] to remove the frequency offset effects. Another proposed approach is the Multiple Symbol Differential Detection (MSDD) demonstrated in [13]. This manuscript aims to provide a survey of the receivers' design in cooperative diversity system so as to alleviate the frequency offset problems.

The manuscript is structured as follows. In Section II, a discussion on overview of the DD transmission scheme is described. In Section III, the concept of MSDD technique in the presence of carrier offset bypassing the channel knowledge incorporate in cooperative diversity is presented. Finally, conclusions and future directions are drawn.

II. THE DOUBLE DIFFERENTIAL TRANSMISSION

The DD transmission approach is one of the techniques employed to mitigate the performance degradation in the presence of frequency offsets. Fig. 2a illustrates the DD encoder block, wherein, $x[n]$ represents the signals in M -Pulse Shift Keying (M -PSK) constellation at time n and D denotes the delay. The transmitted signal $z[n]$ at time n is described by $x[n] \times y[n-1] \times z[n-1]$. In view of a point-to-point wireless communication link with carrier offset, the received signal can be attained as $r[n] = h e^{jw} r[n] + e[n]$ illustrated in Fig. 1(b) where, h represents the channel gain and w is the unknown carrier offset with $e[n]$ denotes the Additive White Gaussian Noise (AWGN).

The decision variable $\hat{x}[n]$ is represented by $p[n]p[n-1]^*$, where $p[n] = r[n]r[n-1]^*$.

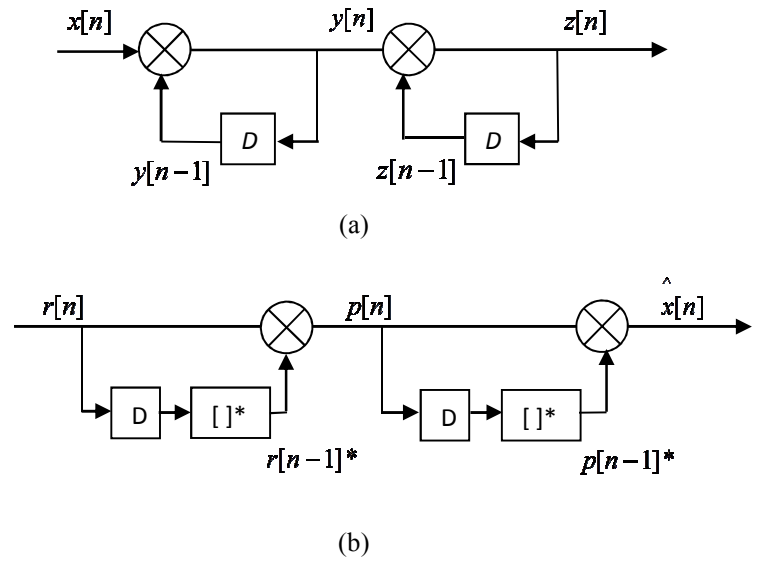


Figure 2: A Block Diagram of Double Differential (DD) (a) Modulation and (b) Demodulation. [14]

By employing this approach, the information detection can be obtained based on the previously received consecutive symbols which bypass the unknown channel coefficients knowledge. This approach is investigated in [14] for the OWRN cooperative diversity system as illustrated in Fig. 3. In this scheme, AF protocol is employed in the presence of random frequency offsets over a block fading Nakagami- m channels which remain static for at least three sequential time intervals. Referring to Fig. 3, the information signal is encoded using DDM (i.e. DD modulation) before being broadcasted to the relay and directly towards the destination in half-duplex transmission manner.

During the first phase, the signal received at the relay is described by:

$$x_{s,r}(n) = \sqrt{P_1} h_{s,r} e^{jw_{s,r}n} z(n) + e_{s,r}(n), \quad n=0, 1, \dots, \quad (1)$$

and the signal received at the destination is written as:

$$x_{s,d}(n) = \sqrt{P_1} h_{s,d} e^{jw_{s,d}n} z(n) + e_{s,d}(n), \quad n=0, 1, \dots, \quad (2)$$

where, P_1 denotes the power transmitted by the source, $h_{s,r}$ and $h_{s,d}$ are the channel gains, $w_{s,r}$ and $w_{s,d}$ are the carrier offsets, and $e_{s,r}$ and $e_{s,d}$ are the Additive White Gaussian Noise (AWGN) between the source-relay and source-destination links respectively.

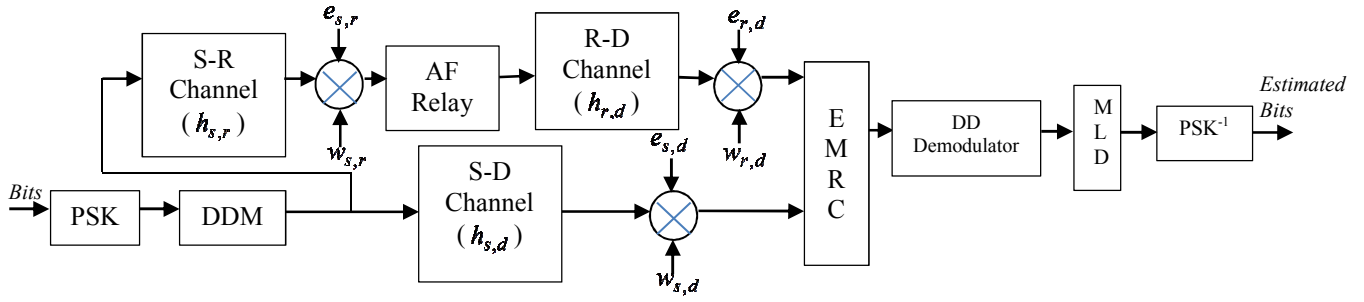


Figure 3 : Block Diagram of the DD transmission. [14]

During the second phase, the source remains idle and the relay amplifies the signal of $x_{s,r}$ in (1) and retransmits the signal towards the destination as:

$$x_{r,d}(l) = \sqrt{P_2} h_{r,d} e^{jw_{r,d}m} x_{s,r}(l) + e_{r,d}(l), \quad l=0, 1, \dots, \quad (3)$$

with l representing the time index that shows the time difference for the first and second phase, P_2 is the amplification factor at the relay, $h_{r,d}$ is the channel gain, $w_{r,d}$ is the frequency offset and $e_{r,d}$ is the AWGN between the relay-destination link.

The relayed and directly transferred signal are combined by utilizing the Maximal Ratio Combining (MRC) [15] scheme and the decision variable $q(n)$ is proposed based on DD demodulator as shown in Fig. 3 as:

$$q(n) = A_1(r_{s,d}[n]r_{s,d}^*[n-1])(r_{s,d}[n-1]r_{s,d}^*[n-2])^* + A_2(r_{r,d}[l]r_{r,d}^*[l-1])(r_{r,d}[l-1]r_{r,d}^*[l-2])^* \quad (4)$$

where, $k = n = l$ and A_1 as well as A_2 are formulated as:

$$A_1 = \frac{1}{(2P_1 |h_{s,d}|^2 + \sigma^2)\sigma^2} \quad (5)$$

$$A_2 = \frac{(P_1\sigma_{s,r}^2 + \sigma^2)^2}{\kappa} \quad (6)$$

where,

$$\begin{aligned} \kappa = & 2P_1P_2^2 |h_{r,d}|^4 |h_{s,r}|^2 \sigma^2 + 2P_1P_2 \\ & \times (P_1\sigma_{s,r}^2 + \sigma^2) |h_{r,d}|^2 |h_{s,r}|^2 \sigma^2 + P_2^2 |h_{r,d}|^4 \sigma^4 + 2P_2 \\ & \times (P_1\sigma_{s,r}^2 + \sigma^2) |h_{r,d}|^2 \sigma^4 + (P_1\sigma_{s,r}^2 + \sigma^2)^2 \sigma^4 \end{aligned} \quad (7)$$

From Eq. 7, in order to eliminate the channel knowledge, the MRC scheme replaces the channel coefficients with the channel variances. Then, the signal is decoded using Maximum Likelihood (ML) approach and can be written as:

$$\hat{x}(n) = \arg \max_{x \in \mathbb{Z}} \text{Re} \{q[k]x^*\} \quad (8)$$

The Bit Error Rate (BER) analysis is provided which revealed that the MRC performed poorer than the ideal MRC in [15]. Thus, instantaneous SNR is suggested and has been proved to perform approximately near to the ideal MRC scheme. In addition, the proposed scheme is able to predict the behaviour of the cooperative system without knowing the explicit channel knowledge and frequency offset. The concept of DD transmission is also implemented in [13–15] respectively. Cano *et al.* in [16] suggested the DD modulation at the relays and simple heuristic detector in a distributive network as shown in Fig. 4.

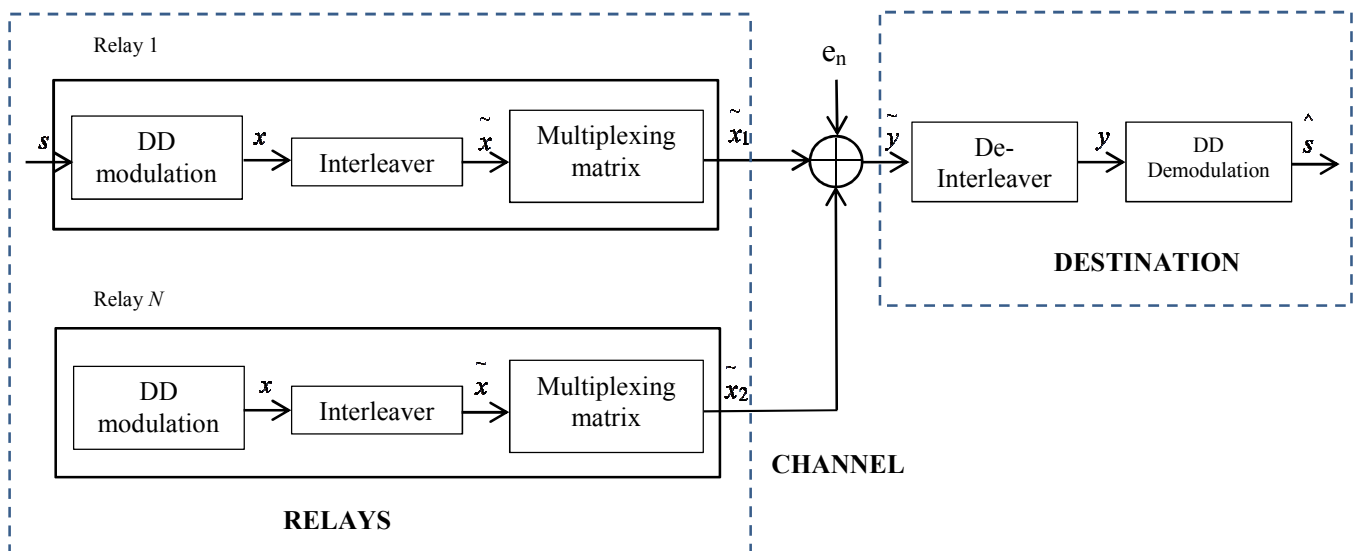


Figure 4 : Proposed DD Transmission System Model [16]

Refer to Fig. 4 the relay receives a block of $K - 2$ symbols of information s . The element of s is then mapped to a unitary diagonal matrix v_k . A recursions approach is applied in the v_k to attain the double-differentially encoded block signals x_k . The symbols x_k are then interleaved to

form \tilde{x} and analysed through a multiplexing matrix. At the destination, the heuristic detector is given by:

$$v_k = \arg \min_{v \in \Xi} \left\| \sum_k^{-1/2} (y_k y_{k-1}^* - v y_{k-1} y_k^*) \right\|^2 \quad (9)$$

The said scheme, however, requires the previous source knowledge and transmission ordering protocol¹ to ensure that each source acknowledge the corresponding relay. Hence, it may be inefficient in an ad-hoc network since the source may join or discontinue from the network [16].

DD transmission in a TWRN based system is further investigated in [17] using AF relaying protocol. Referring to Fig. 1(b), initially the relay initiates and transmits a sequence of signals to both users so that each user can estimate its own channel gain before the transmission process. In Phase 1, both User 1 and User 2 transmits the double-differentially encoded signals via a relay in the presence of Doppler effects. At the relay, the received signals are converted from the carrier frequency to baseband signals. During Phase II, the relay amplifies the complex conjugate of the combined received signal with a multiplication factor. The resulting signal is then converted to its carrier frequency and broadcasted to the desired destination. The signal received encompasses the self-interference as well as the desired signal from its counterpart. By using the DD demodulation, the information symbol of the user can be decoded after the self-interference² cancellation process. The average Symbol Error Rate (SER) shows that the proposed scheme performed better as compared to the OWRN with the assumption that the frequency offset is perfectly estimated and compensated. However, it can be observed that the performance of SER degrades when the pilot symbols are reduced. The need of extra overhead information reduces the bandwidth. Furthermore, the said scheme considers a perfect self-interference cancellation at the destination, wherein, the consideration is impractical in real application [16].

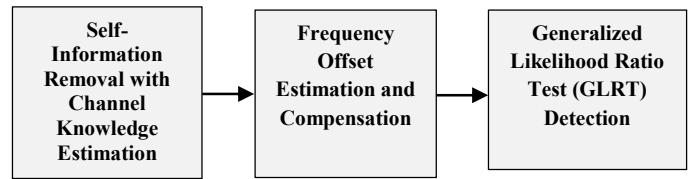
III. MULTIPLE SYMBOL DIFFERENTIAL DETECTION

Multiple Symbol Differential Detection (MSDD) scheme is based on the utilization of more than two consecutive symbols as compared to the symbol-by-symbol based transmission to detect transmitted symbols at the destination.

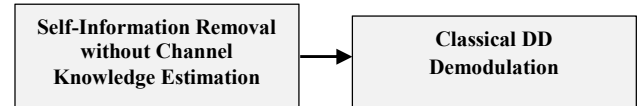
¹Transmission order protocol is the order in determining that which nodes acts as online transmitters, where the source node is the first node in the $N + 1$ network nodes order and ends with the destination node [21].

²Self-interference occurs when an unwanted signal is being received at the destination over a transmission link from the source.

Wu *et al.* [19] presented two schemes, namely Single Differential (SD) and Double Differential (DD) based detections employing AF in flat-fading channel of TWRN system in the presence of frequency offset. For SD, a three-stage approach as shown in Fig. 5(a) is carried out.



(a)



(b)

Figure 5: Multiple Symbol Differential Detection based on
(a) Symbol-By-Symbol based Single Differential Detection
(b) Multiple-Symbol Double Differential Detection

By employing the SD detection as illustrated in Fig. 5, the self-information of the source nodes were removed upon receiving the superimposed information broadcasted from the relay by estimating the channel knowledge (i.e., amplification factor and channel gains). In the following stage, the estimation of frequency offset has been employed by way of transmitting only two training symbols for each information frame. The effect of frequency offset is then compensated following the frequency offset estimation. The received symbols are detected using the Generalized Likelihood Ratio Test (GLRT)-based algorithm³.

In lieu of Double Differential, a two-stage mechanism as depicted in Fig. 5(b) is proposed, wherein, the first stage is the process of single symbol differential detection as well as self-information subtraction bypassing channel estimation. For the second stage, classical DD detector is implemented because of its attractive characteristics that omits the highly-complex frequency offset estimation, compensation as well as the tracking circuitry. The detection mechanism can be executed by employing symbol-by-symbol and multiple-symbol transmission. Multiple-symbol DD detection is attractive due to its ability to recover the desired information without requiring higher SNR, in order to obtain the same average BER performance of the counterpart coherent detection, as compared to DD modulation [19,20] Double differential multiple-symbol detection can be obtained via the single differential basis. The received signal at either source nodes S_1 or S_2 with no additive noise was obtained as:

$$\hat{y}_1(t) = \bar{y}_1(t) \bar{y}_1^*(t-1) = |v|^2 e^{j(\omega_2 - \omega_1)} p_2(t-1) z_2(t) = \hat{h} p_2(t-1) z_2(t) \quad (10)$$

³In a GLRT detection, the likelihood function is estimated or maximized over an unknown parameter (i.e. the channel gains) in order to obtain the decision metrics [22].

wherein, \bar{y}_1 is the received signal after compensation process with $(*)$ representing the conjugate operation; v denotes the channel knowledge; w_1 and w_2 represents unknown carrier offsets; p_2 is the power transmitted by the relay and $z_2(t)$ indicates the recovered symbols. In order to obtain double differential detection, Eq. 10 was iterated as:

$$\hat{y}_1(t-n) = \hat{h}p_2(t-N+2) \prod_{m=0}^{N-n-3} z_2(t-n-m),$$

$$n = 0, 1, 2, \dots, N-2 \quad (11)$$

with N signifying the length of the symbol in the observation. Minimum Least-Square (LS) was applied over the minimization metric of $\hat{h}p_2^*(t-N+2)$ to attain the following decision rule with:

$$\max_{z_2(t), z_2(t-1), \dots, z_2(t-N+3)} \left| \sum_{n=0}^{N-2} y_1(t-n) \dots \prod_{m=0}^{N-n-3} z_2^*(t-n-m) \right|^2 \quad (12)$$

Nevertheless, Eq. 12 showed a high complexity in the computational process of $(MN-3M-N)M^{N-1}/(M-1)^2$. Hence, a fast algorithm is applied so as to reduce the complexity on the order of $N \log_2 N$ [13]. The proposed schemes were analysed in terms of BER and it was observed that the SD detection with multiple symbol transmission degrades when the observation length N increased due to the execution of short training symbols. The limited number of training symbols transmitted results in an inaccurate estimation of the frequency offset. Hence, it was suggested that the single symbol transmission should be adopted in order for the single differential detection to perform in the

presence of frequency offset. Owing to the requirement of additional pilot symbols, the proposed estimator decreases the data rate. Moreover, the frequency offset acquisition and tracking circuitry is correspondingly required which acquires complex hardware implementation. By contrast, multiple symbol detection appears to have the capability to alleviate the loss of SNR resulting from the DD modulation; as compared to the single symbol based detection where the performance significantly improves when N is increased. However, this scheme yields a complex computational process.

Table 1 summarizes the receiver designs depending on a variety of networks, channel estimation as well as frequency offset requirements and pilot-based transmission.

IV. CONCLUSION AND FUTURE WORK

Despite the impressive effect of differential transmission to mitigate the computational complexity of channel estimation, the performance in terms of error probability rate is still facing degradation due to the impact of nodes' mobility on cooperative diversity networks. This survey describes the receiver designs based on DD transmission as well as MSDD in cooperative diversity system for wireless communication in the presence of frequency offset. The survey illustrates that the proposed mechanisms lacks in the solution to dispense the channel coefficient and frequency offset estimation at the same time. These estimations requires the implementation of pilot symbols transmission that reduces the data rates as well as bandwidth because of the rapid channel variations. One prospective solution is the implementation of multi-symbols based transmission without requiring the full or partial knowledge of channel coefficient and frequency offsets.

Table 1: Summary of Receiver Designs in the Survey

| Article | Network | Fading | Relay Protocol | Direct Link | Channel Estimation | Frequency Offset Estimation | Blind/Pilot |
|-----------------------|------------------|-------------------------|----------------|-------------|-------------------------------|---|-------------------------------|
| Bhatnagar et. al [14] | OWRN | Nakagami- m | AF | Yes | Instantaneous SNR is required | No | Partial Pilot |
| Cano et. al [16] | Distributed OWRN | Discrete-Time Low-Pass | - | No | No | No | Pilot |
| Gao et. al [17] | TWRN | Nakagami- m | AF | No | Yes (Self) | Requires frequency offset estimation and compensation | Pilot (Transmission Overhead) |
| Wu et. al [19] | TWRN | Frequency Flat Rayleigh | AF | No | Yes (SD) No (DD) | Requires frequency offset estimation and compensation | Both |

REFERENCES

- [1] A. Sendonaris, E. Erkip and B. Aazhang. (2003). User cooperation diversity – Part I: System description. *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938.
- [2] A. Nosratinia, T. E. Hunter, and A. Hedayat. (2004). Cooperative communication in wireless networks. *Commun. Mag. IEEE*, vol. 42, no. 10, pp. 74–80.
- [3] J. Liao, F. Wang, D. Yao, and M. Wang. "Which is Better: One-way or Two-way Relaying with an Amplify-and-Forward Relay?" in *Wireless Communications and Networking Conference (WCNC)*, 2014 IEEE (pp. 1087-1092). IEEE.
- [4] J. Tian, Q. Zhang, and F. Yu. Non-coherent detection for two-way AF cooperative communications in fast rayleigh fading channels. (2011). *IEEE Trans. Commun.*, vol. 59, no. 10, pp. 2753–2762.
- [5] J. N. Laneman, D. N. C. Tse, and G. W. Wornell. (2004). Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080.

- [6] T. Issariyakul and V. Krishnamurthy. (2009). Amplify-and-Forward Cooperative Diversity Wireless Networks: Model, Analysis, and Monotonicity Properties. *IEEE/ACM Trans. Netw.*, vol. 17, no. 1, pp. 225–238.
- [7] T. Himsoon, W. Su, and K. J. R. Liu. (2005). Differential transmission for amplify-and-forward cooperative communications. *IEEE Signal Process. Lett.*, vol. 12, no. 9, pp. 597–600.
- [8] Q. Zhao and H. Li. (2005). Performance of differential modulation with wireless relays in Rayleigh fading channels. *IEEE Commun. Lett.*, vol. 9, no. 4, pp. 343–345.
- [9] T. Himsoon, W. P. Siriwoongpairat, W. Su, and K. J. R. Liu. (2008). Differential modulations for multinode cooperative communications. *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2941–2956.
- [10] M. El-Hajjar, L. Hanzo, C. Versus, and N. Communication. (2010). Dispensing with channel estimation.... *IEEE Veh. Technol. Mag.*, no. June, pp. 42–48.
- [11] J. Liu, P. Stoica, M. Simon, and J. Li, “Single differential modulation and detection for MPSK in the presence of unknown frequency offset,” in *Conf. Rec. - Asilomar Conf. Signals, Syst. Comput.*, pp. 1440–1444, 2006.
- [12] M. K. Simon and D. Divsalar. (1992). On the implementation and performance of single and double differential detection schemes. *IEEE Trans. Commun.*, vol. 40, no. 2, pp. 278–291.
- [13] D. Divsalar and M. K. Simon. (1990). Multiple-symbol differential detection of MPSK. *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300–308.
- [14] M. R. Bhatnagar, A. Hjrøngnes, and L. Song. (2008). Cooperative communications over flat fading channels with carrier offsets: A double-differential modulation approach. *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 531786, pp. 1–11.
- [15] D. G. Brennan. (2003). Linear diversity combining techniques. *Proc. IEEE*, vol. 91, no. 2, pp. 331–356.
- [16] A. Cano, E. Morgado, A. Caama, and F. J. Ramos. (2007). Distributed Double-Differential Modulation for Cooperative Communications under CFO. pp. 3437–3441.
- [17] Z. Gao, L. Sun, Y. Wang, and X. Liao. (2014). Double Differential Transmission for Amplify-and-Forward Two-Way Relay Systems. *Communications Letters, IEEE*, vol. 18, no. 10, pp. 1839–1842.
- [18] M. R. Bhatnagar, A. Hjørøngnes, and L. Song. (2008). Double-differential decode-and-forward cooperative communications over nakagami-m channels with carrier offsets. *Sarnoff Symp. IEEE*, pp. 1–5, 2008.
- [19] Z. Wu, G. Li, and T. Wang, “Differential modulation for amplify-and-forward two-way relaying with carrier offsets,” in *IEEE Int. Conf.*, pp. 4501–4506, 2014.
- [20] M. Simon, J. Liu, P. Stoica, and J. Li. (2004). Multiple-Symbol Double-Differential Detection Based on Least-Squares and Generalized-Likelihood Ratio Criteria. *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 46–49.
- [21] S. C. Draper, L. Liu, A. F. Molisch, and J. S. Yedidia. (2011). Cooperative transmission for wireless networks using mutual-information accumulation. *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5151–5162.
- [22] M. R. Souryal. (2010). Non-coherent amplify-and-forward generalized likelihood ratio test receiver. *IEEE Trans. Wirel. Commun.*, vol. 9, no. 7, pp. 2320–2327.

Measuring and Predicting the Impacts of Business Rules Using Fuzzy Logic

Mohamed Osman Hegazi

Department of Computer Science, College of Computer Engineering and Science, Prince Sattam University , Saudi Arabia
Department of Computer Science, College of Computer Science and IT , Alzaiem Alazhari University Sudan

Abstract- The relations between the constituent elements of the predominant activities may cannot be measured using the accurate measurement, the reason may be on the conflict or the effect of such rules with other factor, accordingly uncertainty measurement is needed. This paper presents a measurement and predictive fuzzy model that can be fit to work on such rules. The proposed model measures and predicting the impacts of such rules using fuzzy logic, it designed based on the concept of fuzzy logic, fuzzy set and the production rule. The model transforms if-then business rule to weighted fuzzy production rule, and then used this production rule for predicting and measuring the impact of the business rule. The model is tested using real data and provide considerable results.

Keyword: Fuzzy logic, Fuzzy set, Production rule, Business rule

I. INTRODUCTION

Fuzzy logic and fuzzy set is a huge field that includes fuzzy measure, fuzzy topology, fuzzy algebra, fuzzy analysis, etc. Only a small portion of fuzzy set theory has been applied to business problems. A fuzzy set member takes values in the interval $[0,1]$. Therefore, a fuzzy set is a generalization of a classical set by allowing the membership function to take values in the interval $[0,1]$ instead of just membership [5].

We mean by business rule any rule that dealing with the relations between the constituent elements of the predominant activities, whether that activity is in the institutions of business administration, economics, or scientific and engineering applications. For examples the relation between the advertising and sales, the relation between power and performance of the hardware in engineering sectors, the relation between money supply and inflation ... etc. These kinds of rules may cannot be measured using the accurate measurement, the reason may be on the conflict or the effect of such rules with other factor, accordingly uncertainty measurement is needed. This paper presents a measurement and predictive model that can be fit to work on such rules by using the concepts of fuzzy logic and fuzzy set.

There are many studies that dealt with the fuzzy logic in the field of business activities, or in formalizing fuzzy models such as:

The study of Chen, Shyi-Ming, and Ze-Jin Chen [3], they provide weighted interpolative fuzzy reasoning model for dealing with regression problems, their model uses the representative values in the fuzzy set, calculate the piecewise fuzzy between any two characteristic points, calculate the weights and finally, calculate the degree of membership of each obtained characteristic point of the fuzzy interpolative result.

In [1] a fuzzy rule based system is developed to solve weather event prediction for Lahore. They provide a procedure for designing an automated rule generator based on rule base reasoning.

In [2] fuzzy interpolative reasoning method for sparse fuzzy-rule-based systems is proposed. They used the weighted average to infer the fuzzy interpolative reasoning results

In [6] a fuzzy production rule based expert system is developed using PROLOG language to help physicians. The system used fuzzy vector as input and produces a fuzzy set as output.

In [4] an evaluation and generation of nonlinear integral tool in the fuzzy production rule reasoning. They generate application of similarity-based reasoning.

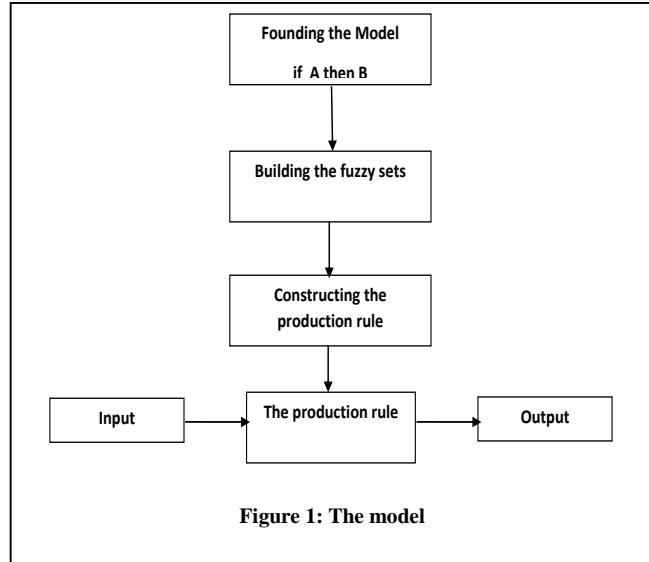
In [7] a fuzzy expert system is proposed to evaluate the performance degree of intellectual capital. Their expert system formalizing the Fuzzy Sets and generate the fuzzy Knowledge Acquisition. The proposed expert system is designed to be able to understand the enterprise situation and then performance degree of intellectual capital

The rest of the paper is organized as the following:
Section II present the suggested model, in section III the

model is tested (the experiment is presented), and finally section IV concludes the paper.

II. THE SUGGESTED MODEL

The suggested model based on the concepts of fuzzy logic and fuzzy set. The model is developed according to the step presented in figure 1.



A. Determine the Business Rule (founding the model):

The model is based on the “if then rule” (equation 1)

$$\text{if } A \text{ then } B \quad \text{or} \quad A \Rightarrow B \quad (1) \text{-a}$$

or;

$$\text{if } \text{if } A \text{ (to some degree } x \text{) then } B \text{ (to some degree)} \quad (1) \text{-b}$$

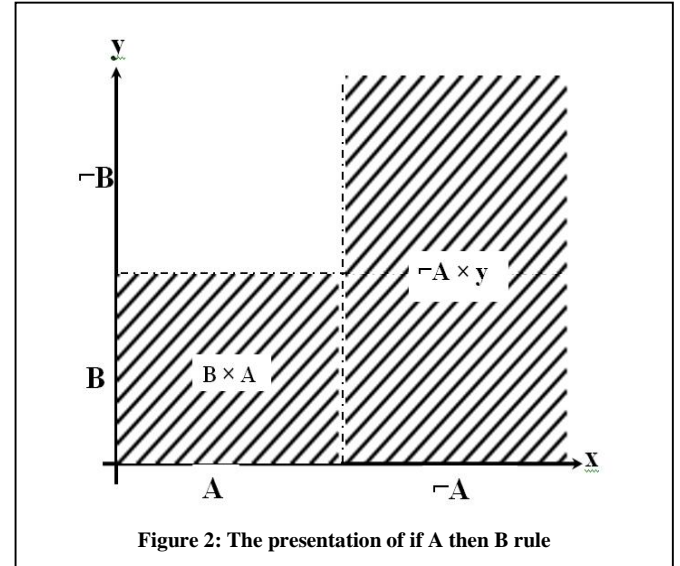
This rule (equation 1 –a, and b) logically is equal to $(A \text{ and } B) \cup \neg A$ (equation-2) that because if A and B means:

- 1- if A true then B will be true
- 2- if A not true then B may be true or not true, for example : If increased spending on advertising then revenue will increase. We can conclude from this example it will be acceptable if revenue increase even spending on advertising not increase. Accordingly, “if A then B” rule accept all cases of B (True or False).

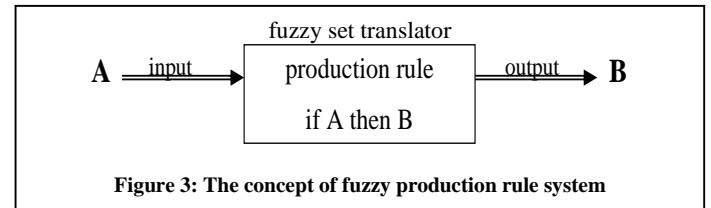
Therefore the rule if A then B can be presented logically as follow:

$$\text{if } A \text{ then } B = (A \times B) \cup (\neg A \times Y) \quad (2)$$

As it shown in figure (2) the rule "if A then B" is presented on the shaded two parts the part where A and B are true (the intersection of $B \times A$) and the intersection between $\neg A$ and all Y ($\neg A$ or $(B$ or $\neg B)$)



The rule in equation 2 can be used as a production rule in developing our fuzzy model as it shown in the figure (3)



B. Building the fuzzy sets:

For every data items we build a fuzzy set using the following algorithm:

- 1- Determine the normal set of the elements $\{i_1, i_2, i_3, \dots, i_n\}$
- 2- Find the largest value (G_i) of the set elements ($G_i = \max\{i_1, i_2, i_3, \dots, i_n\}$).
- 3- Build the fuzzy set using the following formula:

$$\text{Fuzzy set} = \{i_1 / G_i, i_2 / G_i, i_3 / G_i, \dots, i_n / G_i\} \quad (3)$$

Hence:

$$\begin{aligned} i_1, i_2, i_3, \dots, i_n &\rightarrow \text{the set elements (set members)} \\ G_i &\rightarrow \text{the largest value (Max}\{i_1, i_2, i_3, \dots, i_n\}) \end{aligned}$$

$i_1/G_i, i_1/G_i, i_1/G_i, \dots, i_1/G_i \rightarrow$ the fuzzy set element (fuzzy set members)

For example if we want to build a fuzzy set for the inflation rate (table 1):

- 1- The normal set is:
{69,130.3, 46.5, 17.7, 16.1, 8.1}
- 2- The largest value G_i is:

$$G_i = 130.3$$

- 3- Build the fuzzy set using the following formula:
Fuzzy set = { {69/130.3, 130.3/130.3, 46.5/130.3, 17.7/130.3, 16.1/130.3, 8.1/130.3} }
Fuzzy set = {0.53, 1.00, 0.36, 0.14, 0.12, 0.06}

We can see that the final fuzzy set, in this example, providing the membership degree of each element instead of only the membership as it is in the normal set. This is one of benefit of using the fuzzy set.

C. Constructing the product rule:

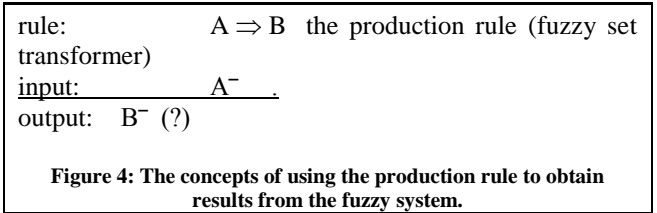
In this stage we used the results of stage 2 above (the fuzzy sets) to construct the production rule if then (equation 2) as follow:

- 1- Determine the component of the rule (A,B and the relation between them) this rule is determine according to the business rule (*If increased spending on advertising then revenue will increase*)
- 2- Determine the fuzzy set that represents A in our production rule (for example the advertising cost).
- 3- Determine the fuzzy set that represents B in our production rule (for example the revenue)
- 4- Applying the rule (equation 2). Noting that the mathematical operation is done using the logical concept i.e. we used the logical AND instead of the multiplication and the logical OR instead of addition or union operation.

The result of this step is a matrix that represents the fuzzy set transformer.

D. Make used of the product rule:

Finally the result of stage 3 above can be used for developing results that can provide the measurement or the prediction of the impact of the business rule. This is done by entering an input vector (fuzzy set) into the production rule – fuzzy rule transformer- the output will be the needed result (figure 3)



III. TESTING THE MODEL(THE EXPERIMENT)

Let us used the data provided by Sudan Central Bank report (table 1). In this data the Government funding from the Bank of Sudan rate for the years starting from 1995 to2000 is presented in column 2 (this data is real data, it is about the rate of government funding in respect to the total banking finance) and the rate of the inflation is presented in column 2

TABLE 1: THE RELATION BETWEEN GOVERNMENT FUNDING AND THE RATE OF INFLATION (for the period 1995-2000)

| years | Government funding from the Bank of Sudan rate (billion pound) | Inflation rate |
|-------|--|----------------|
| 1995 | 26.5 | 69 |
| 1996 | 32.4 | 130.3 |
| 1997 | 7.9 | 46.5 |
| 1998 | 9 | 17.7 |
| 1999 | 12 | 16.1 |
| 2000 | 8 | 8.1 |

Source : Bank of Sudan report (<http://www.cbos.gov.sd/node/211>)

- 1- If there is business rule for the relation between the government funding and the inflation. If the rule is: If there the government takes fund from the central bank (A) then there will be growth in the inflation rate (B). This rule will be equal to the following fuzzy logical rule:

$$\text{if } A \text{ then } B = (A \times B) \cup (\neg A \times Y)$$

equation (2) above

where A: Government fund, and B: inflation rate

- 2- Building the fuzzy sets:
 - Building the fuzzy set for the government fund (A):
 - The normal set for government fund A= {26.5, 32.4, 7.9, 9, 12, 8}
 - The largest value G_i is: $G_i = 32.4$
 - The fuzzy set = {26.5/32.4, 32.4/32.4, 7.9/32.4, 9/32.4, 12/32.4, 8/32.4}
 - A = {0.82, 1.00, 0.24, 0.28, 0.37, 0.25}
 - Building the fuzzy set for the inflation rate (B):
 - The normal set for the inflation rate B= {69,130.3, 46.5, 17.7, 16.1, 8.1}
 - The largest value G_i is: $G_i = 130.3$

- The fuzzy set = {69/130.3, 130.3/130.3, 46.5/130.3, 17.7/130.3, 16.1/130.3, 8.1/130.3}

$$B = \{0.53, 1.00, 0.36, 0.14, 0.12, 0.06\}$$

- Calculate the negation of A ($\neg A$):
- Not A = $1-A$

$$\neg A = \{1-0.82, 1-1, 1-0.24, 1-0.28, 1-0.37, 1-0.25\}$$

$$= \{0.18, 0, 0.76, 0.72, 0.63, 0.75\}$$

$$= \begin{pmatrix} 0.82 \\ 1 \\ 0.24 \\ 0.28 \\ 0.37 \\ 0.25 \end{pmatrix} \circ (0.53 \ 1 \ 0.36 \ 0.14 \ 0.12 \ 0.06) \cup \begin{pmatrix} 0.18 \\ 0 \\ 0.76 \\ 0.72 \\ 0.63 \\ 0.75 \end{pmatrix} \circ (1 \ 1 \ 1 \ 1 \ 1 \ 1)$$

$$= \begin{pmatrix} 0.53 & 0.82 & 0.36 & 0.14 & 0.12 & 0.06 \\ 0.53 & 1 & 0.36 & 0.14 & 0.12 & 0.06 \\ 0.24 & 0.24 & 0.24 & 0.14 & 0.12 & 0.06 \\ 0.28 & 0.28 & 0.28 & 0.14 & 0.12 & 0.06 \\ 0.37 & 0.37 & 0.36 & 0.14 & 0.12 & 0.06 \\ 0.25 & 0.25 & 0.25 & 0.14 & 0.12 & 0.06 \end{pmatrix} \cup \begin{pmatrix} 0.18 & 0.18 & 0.18 & 0.18 & 0.18 & 0.18 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0.76 & 0.76 & 0.76 & 0.76 & 0.76 & 0.76 \\ 0.72 & 0.72 & 0.72 & 0.72 & 0.72 & 0.72 \\ 0.63 & 0.63 & 0.63 & 0.63 & 0.63 & 0.63 \\ 0.75 & 0.75 & 0.75 & 0.75 & 0.75 & 0.75 \end{pmatrix}$$

$$= \begin{pmatrix} 0.53 & 0.82 & 0.36 & 0.18 & 0.18 & 0.18 \\ 0.53 & 1 & 0.36 & 0.14 & 0.12 & 0.06 \\ 0.76 & 0.76 & 0.76 & 0.76 & 0.76 & 0.76 \\ 0.72 & 0.72 & 0.72 & 0.72 & 0.72 & 0.72 \\ 0.37 & 0.37 & 0.36 & 0.14 & 0.12 & 0.06 \\ 0.75 & 0.75 & 0.75 & 0.75 & 0.75 & 0.75 \end{pmatrix} \quad (3)$$

This matrix present the production rule for this example (the fuzzy set transformer)

Note: the operation for handling this equation is done logically (AND instead of multiplication, OR instead of addition and union)

4- To make used of the production rule we can apply the concepts presented in figure 3 and 4 as follow:

- First: Selecting the input set (the set that will be used as input for measuring or predicting the system): for example if we want to see what will happen if the government funding is increase, in this case the input set can be obtain by scaling the funding fuzzy set (A- above) by power 2:

- The Y vector = {1, 1, 1, 1, 1, 1}

3- Construct the production rule: In this stage we will generate the production rule using the above fuzzy set as follow:

- The production rule:

$$\text{if } A \text{ then } B = (A \times B) \cup (\neg A \times Y)$$

- Building the production rule:

Increasing the government fund by powering A to the power 2

$$A^2 = \{0.82^2, 1.00^2, 0.24^2, 0.28^2, 0.37^2, 0.25^2\} \\ = \{0.67, 1.00, 0.08, 0.14, 0.06\}$$

(Note: we can used different power example 3, half (1/2) .. etc according to what we need to masseur)

- Second : use the new set as an input for the system by multiplied the set with production rule matrix (3) as follow:

$$= \begin{pmatrix} 0.53 & 0.82 & 0.36 & 0.18 & 0.18 & 0.18 \\ 0.53 & 1 & 0.36 & 0.14 & 0.12 & 0.06 \\ 0.76 & 0.76 & 0.76 & 0.76 & 0.76 & 0.76 \\ 0.72 & 0.72 & 0.72 & 0.72 & 0.72 & 0.72 \\ 0.37 & 0.37 & 0.36 & 0.14 & 0.12 & 0.06 \\ 0.75 & 0.75 & 0.75 & 0.75 & 0.75 & 0.75 \end{pmatrix} \times \begin{pmatrix} 0.67 \\ 1.00 \\ 0.06 \\ 0.08 \\ 0.14 \\ 0.06 \end{pmatrix}$$

- c. Last: the output will be a new vector (new fuzzy set) that represents the result or the answer for the question what will happen if the government fund from bank is increase. In this example the output is =

$$\{0.82, 1.00, 0.76, 0.72, 0.37, 0.75\}$$

which equal to:

$$\{0.82*130.3, 1*130.3, 0.76*130.3, 0.72*130.3, 0.37*130.3, 0.75*130.3\}$$

i.e. the actual inflation will be = {106.8, 130.3, 99.0, 93.8, 48.2, 97.7}

This represents the changes on the inflation rate according to the changed on the government fund. If we compare this set with the old one (B) we can see how much the increasing in the government affected the inflation rate.

This model can be used in many ways, for example, if there is a business rule that determine the limited acceptable amount of funding of the government from the central bank (say 10 billion) then the set of over rate fund will be the element that belong to this set with respect to the largest value of these elements as follow:

$$= \{26.5/32.4, 32.4/32.4, 0, 0, 12/32.4, 0\}$$

$$= \{0.82, 1.00, 0, 0, 0.37, 0\}$$

$$\begin{pmatrix} 0.82 \\ 1.00 \\ 0 \\ 0 \\ 0.37 \\ 0 \end{pmatrix} \circ \begin{pmatrix} 0.53 & 1.00 & 0.36 & 0 & 0 & 0 \end{pmatrix} \circ \begin{pmatrix} 0.18 \\ 0 \\ 1 \\ 1 \\ 0.63 \\ 0 \end{pmatrix} \circ \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0.53 & 0.82 & 0.36 & 0.00 & 0.00 & 0.00 \\ 0.53 & 1 & 0.36 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.37 & 0.37 & 0.36 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \end{pmatrix} \cup \begin{pmatrix} 0.18 & 0.18 & 0.18 & 0.18 & 0.18 & 0.18 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0.63 & 0.63 & 0.63 & 0.63 & 0.63 & 0.63 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0.53 & 0.82 & 0.36 & 0.18 & 0.18 & 0.18 \\ 0.53 & 1 & 0.36 & 0.00 & 0.00 & 0.00 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0.63 & 0.63 & 0.63 & 0.63 & 0.63 & 0.63 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \end{pmatrix}$$

This matrix present the production rule for this example (the fuzzy set transformer)

As we can see the third, fourth and the sixth element are not belonging to this set, while the first element is %82 belong to this set, the second element is 100% belong to this set and fifth element is %37 belong to this set. This kinds of set is easy to be generated according to what measurement is needed, also this kind of set can be used to generate production rule according to any business rule, as follow:

The business rule:

If there the government fund that taken from the central bank is more than 10 billion (A) then there will be high rate of inflation (more than 30%) (B)

$$A: G_i = 32.4$$

$$\text{Then } A = \{26.5/32.4, 32.4/32.4, 0, 0, 12/32.4, 0\}$$

$$= \{0.82, 1.00, 0, 0, 0.37, 0\}$$

$$B: G_i = 130.3$$

$$\text{Then } B = \{69/130.3, 130.3/130.3, 46.5/130.3, 0, 0, 0\}$$

$$= \{0.53, 1.00, 0.36, 0, 0, 0\}$$

$$\text{if } A \text{ then } B = (A \times B) \cup (\neg A \times Y)$$

According to concepts presented in figure 3 and 4, this production rule can then be used to produce any results regarding the basic business rule (if the fund exceed than 10 billion), as we did in the previous example.

IV. CONCLUSION

In this paper, we proposed a model that can be use in developing a weighted fuzzy production rule for if-then business rule. Then the model provides the way of using this production rule in measuring and predicting the impact of the business rule. The proposed model is tested in real data and provide considerable results. The model provides a flexibility in changing or scaling the rules, because it based on weighted fuzzy sets, and according to the weighted fuzzy set concepts, it can easy to modify the weight or the membership in the fuzzy set, so any change in rules is just changing in the set weight or membership.

This model can be use in decision support, and it can be enhance and implemented as a computer system or packed.

- [6] Yue, Zhang, et al. "A fuzzy production rule based expert system." *Fuzzy sets and systems* 44.3 (1991): 391-403.
- [7] Zarandi, Mohammad Hossein Fazel, Neda Mohammadhasan, and Susan Bastani. "A fuzzy rule-based expert system for evaluating intellectual capital." *Advances in Fuzzy Systems* 2012 (2012): 7.



Mohamed O. Hegazi has got his Ph.D. in Distributed Database from Sudan University Of Science & Technology, Khartoum Sudan in 2004. He has worked as an associate professor in Alzaiem Alazhari University, Khartoum Sudan for 4 years, he then joined the College of Computer Engineering and Sciences of Prince Sattam Bin Abdulaziz University since 2013 as an associate professor of computer science. His research interests are database, data mining and computer applications.

References

- [1] Awan, Malik Shahzad Kaleem, and Mian Muhammad Awais. "Predicting weather events using fuzzy rule based system." *Applied Soft Computing* 11.1 (2011): 56-63.
- [2] Chang, Yu-Chuan, Shyi-Ming Chen, and Churn-Jung Liau. "Fuzzy interpolative reasoning for sparse fuzzy-rule-based systems based on the areas of fuzzy sets." *Fuzzy Systems, IEEE Transactions on* 16.5 (2008): 1285-1301.
- [3] Chen, Shyi-Ming, and Ze-Jin Chen. "Weighted fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on piecewise fuzzy entropies of fuzzy sets." *Information Sciences* 329 (2016): 503-523. Online 2015
- [4] Yeung, D.S.; Ha, M.H.; Wang, X.Z., "Weighted fuzzy production rule reasoning with interactive propositions," in *IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th* , vol., no., pp.2618-2623 vol.5, 25-28 July 2001
- [5] Zhang, Huaguang, and Derong Liu. *Fuzzy modeling and fuzzy control*. Springer Science & Business Media, 2006.

Blind User Visualization and Interaction in Touch Screen: A Designer Perspective

Abdelrahman H. Hussein

College of Computer Science and Engineering
University of Hail
Hail, Saudi Arabia

Abstract— In this paper, we describe how blind students views external system using an image map as a case study. We proposed two interaction techniques which allow blind students to discover different parts of the system by interacting with a touch screen interface. An evaluation of our techniques reveals that 1) building an internal visualization, interaction technique and metadata of the external structure plays a vital role 2) blind students prefer the system to be designed based upon their behavioural model to easily access and build the visualization on their own and 3) to be an exact replica of visualization, the metadata of the internal visualization is to be provided either through audio cue or domain expert (educator). Participants who used touch screen are novice users, but they have enough experience on desktop computers using screen readers. The implications of this study to answer the research questions are discussed.

Keywords- Blind; Visualization; Touch Screen; Accessibility; Usability; Image Map.

I. INTRODUCTION

Interaction with touch screen is a complex process for blind users. Understanding the internal visualization of cognitive activities is an important process for a designer to design external visualization. Scaife and Rogers (1996) pointed out that more emphasis should be provided by “the cognitive activities when interaction takes place with external visualization, the properties of the external and internal structures and their benefits with different visual representation” [14].

Under these circumstances, it is essential to understand internal visualization, their relation with external visualization and how physical activities help to bind these two visualizations. Without prior knowledge of these abstract concepts, it will be intricate for a novice designer to build external visualization. As a result, it is significant to make cognition a research agenda for building information visualization [8].

To address this research agenda, we propose the following research questions in accordance with mobile learning (m-learning): 1) What is meant by “internal visualization” for a blind user in terms of image map?; 2) For a given external visualization, how it can be related to internal visualization?; and, 3) How do physical activities relate external visualization

with internal visualization? This paper is derived to discuss the answer for the above mentioned research questions.

II. INTERNAL VISUALIZATION

The researchers suggest that blind people have an internal representation for the information they hear. Most current research is on geometrical shapes and mathematical symbols, and has not touched the core representation of “abstract” information. In particular, the information visualization does not vary much within the sighted user. On the other hand, visualization varies widely among blind users depending on the description of narrator, comprehension and prior experience. Analysing the mental model in the literature is the first step to investigate internal representation which may yield an effective theoretical concept.

1.1 Mental Models of Blind

A review of the relevant literature suggests that the mental model is an internal representation of real world phenomena, which is composed of many small-scale representations. Craik (1943) claims that mental model predicts future action by constructing small-scale models, in his book “The Nature of Explanation” [3]. The Rouse and Morris (1986) argue that the application of a specific mental model is necessary [11]. The emergence of mental models has led to a decline in the theoretical concept to HCI [12].

The literature also suggests that there are two potential and influential mental models: the Norman model and the Johnson-Laird model (see Fig. 1).

Both models suggest that there are four agents which influence the interaction between the system and user: 1) The target system which refers to the device the user is intended to use.; 2) A conceptual model developed by designers in which the system is developed; 3) The user model developed by the user through interaction with the system. This user model continues to update whenever the user is exposed to the system; and 4) The model for the user model, as understood by the scientist (Psychologist or Usability Expert) [10].

Norman revealed that mental models are not usually accurate and are highly “volatile”. Nevertheless, they are

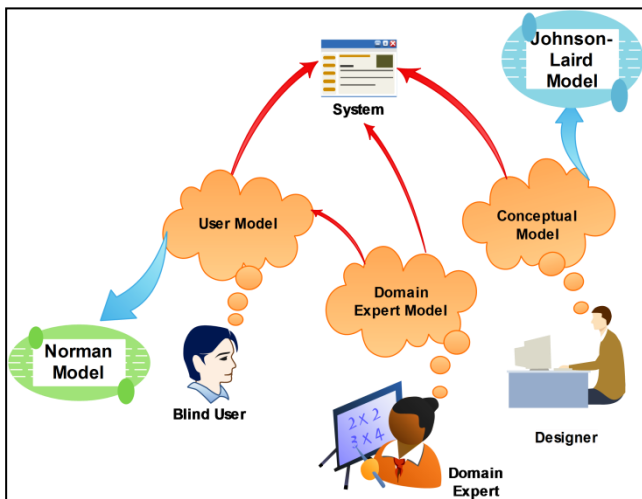


Fig. 1. Mental Model

“runnable” to serve certain rationale. Although there are many lapses, it provides prognostic and expounding power in understanding the interaction.

The Norman model focuses on the user model while Johnson-Laird (1983) focuses on the conceptual model of the system [6]. To put it more simply, Norman emphasizes the “behavioural” aspects of the user towards the system whereas Johnson-Laird defines the “Structural” aspect of the system.

Johnson-Laird views mental models as preserved entities and analogous of what they represent in the external visualization. The user can manipulate the system and configure his mental model with presupposing mental logic and formal rules. However, Norman believes that there are chances the user can pretend to behave with the system. As a consequence, users are forced to give a reason through verbal protocol, although they do not have one.

The Johnson-Laird and Norman models vary widely based on the emphasis and the nature of the problems being investigated. For instance, consider the premises “Tiger eats deer” and “Tiger eats rabbit”. The user can conceptualize two different models where the relations between the entities are present as below:

Table 1 – User mental model and their premises

| Mental Model | Premises1 | Premises2 | Premises3 |
|--------------|-----------|-----------|-----------|
| M1 | Tiger | Deer | Rabbit |
| M2 | Tiger | Rabbit | Deer |

Both models are coherent with the premises; however, the user cannot predict whether deer, eats rabbit or vice versa by merely exploring their mental models. The user cannot apply logical rules to infer, since it does not store premises based on logical predicates such as eat (Tiger, Deer).

1.2 Significance and Application of Information Visualization

To understand the problem related to interface design, it is imperative for designers to understand: 1) how the blind user imagines the external visualization (Section 2.2.1); 2) which tools are used to stimulate these images (Section 2.2.2); 3) the role of data in image formation (Section 2.2.2); and, 4) how the data are mapped together to form an image (Section 2.2.3). To address these questions, we will discuss this in the following section.

1.2.1 Image or Model

Although the items are scattered throughout the web page, the blind user considers all the items in a web page to be a vertical list [9]. The desktop screen reader processes all pages, and produces output in a sequential order that can be navigated by tab or up/down keys.

Currently, touch screen technologies such as the iPhone and Android implement a static layout for the interaction [4]. With reference to this, two dimensional pages are collapsed to single dimensional to form a single horizontal list that contains a large set of items. It is burdened for the user to memorize the sequence of interest items. Furthermore, the relative position of item such as on the top or on the bottom is lost [1]. Although it has limitations, it is considered to be better than nothing.

Applying the interface design format such as size, position and color to the mental model, Johnson-Laird contends that mental models are essentially spatial representations and are more abstract. This inspection is in distinction with mental imagery. He emphasized that both the mental imagery and mental model can be used in logical analysis.

Such characterization makes it challenging to apply to the interface design, especially for blind users for two reasons: 1) The imagery varies widely between the sighted and blind user; and, 2) the designer cannot understand the imagery of the blind user. For instance, the designer can understand how the sighted user can imagine the size and shape of the checklist, but it is difficult to predict a blind user’s imagination about the same fact.

In interface design, the blind user can understand the spatial layout of the screen based on the training. Instead, it seems unreasonable to provide audio cues related the color and length of the widget. Our own experience of developing courseware for the blind user indicates that a blind user has to be provided with cues about the position of the widget in the touch screen device. By continuous exploration, the blind user becomes familiar with the position of the widget [5].

1.2.2 Mental Simulation

The mental model is abstract and cannot be evaluated directly. The review of literature suggests that there are two types to simulate the mental models. In the first type, training the user to use the system and simulate the mental model. In

the second type, the user is exposed to complex system without proper training and simulates the mental model [16]. Two factors are used to simulate the mental models: tools and data.

Tools: Different traditional tools can be applied to simulate the mental model for blind users [13]: 1) **Task-based scenario:** It requires users to perform a series of tasks to achieve the target. Generally, task acquisition time, number of clicks and error rate of the task performance are gathered. The user is prompted to think aloud during the experiment. It helps the investigator to extract the reason for performed errors; 2) **Verbal and Hands on Scenario:** It includes a task-based scenario, but it requires the user to respond to a query pertaining to the task performed with the system. Finally, toward the end of the experiment, the user has to explain the system; 3) **Tutoring Scenario:** The user is required to tutor about the system to another person. It is suitable if the user has enough experience to handle the system. Furthermore, the user has to depend on the learner. However, it encourages the users to articulate their knowledge; and 4) **Exploration Scenario:** Under this scenario, the user is involved with another person and investigates the application. In this way, the users were communicative and more involved. The main impediment to this type is a highly experienced user who dominates the session and less skilled users who will be controlled due to deficiencies in exposure. Additionally, it consumes much time to gather the data which is extracted through recordings.

Data: Mental models do not possess any data. The information about the external visualization is stored in the form of data which we termed as external data. The external data is used to simulate the mental model. The external data may be constructive data or passive data. Constructive data is data that is transferred from external visualization of the user in the form of audio cues to stimulate the next course of action. On the other hand, passive data is the acknowledgment sent by the system to users in the form of haptic feedback. Along these lines, data are transmitted to humans through the hands and ears of the human body. The transmission of data from external visualization of the user simulate the mental model by analyzing the data, selecting the task and choosing appropriate interaction techniques to accomplish the task. For instance, data named "Enter your name" invokes the user to search for the required information (name) in the internal data, select the data, and deliver the data using a voice synthesizer interaction technique. Finally, item level information is stored in the form of data. The information may be individual, such as "Click here" or may be aggregate information such as "error code".

1.2.3 Mental Map

The mental map is the internal representation of data in the mental model, the available data are mapped to each based on interaction with the system. The mental map is primarily related to item-level information. According to the schematic and semantic level, data are aggregated and the relevant task

is performed to construct and simulate the mental model for different visualizations to get a feel of interface design. Thus, the mental model is more abstract than the mental maps.

III. DYNAMICS BETWEEN MENTAL MODEL AND EXTERNAL VISUALIZATION

The specification for e-assessment has to coexist with standard HCI usability (content usable, easy to use, ease to learn, intuitive). The usability requirement for both blind and sighted users is needed on assessment resources for each assessment type and assessment method.

As discussed earlier, the blind user who uses the latest technologies, such as the iPhone and Android, views the external visualization as a horizontal list. Whatever the layout design is, a blind user navigates in the form of a queue keeping the layout static. The navigation is achieved through flung gesture, which supports both back and forward directions.

In Fig. 2 the dynamic interaction that we devised, the blind user views the external visualization as dots in Braille located at fixed locations in the touch screen interaction. For each blind user, each dot acts as a stack in which items pile on top of each other. In this manner, the blind user can navigate through the items in the stack by click gesture. Since many stacks can be placed on the screen, many items can be accommodated which can be reached easily as compared to a static layout. The only predicament the blind user faces will be identification of the required stack. Navigation is supported by external data, such as audio feedback.

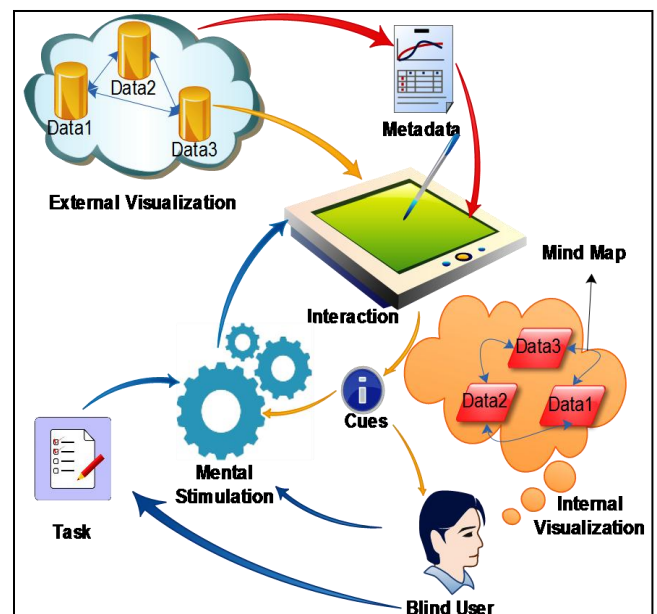


Fig. 2 – Dynamics between the mental models and external visualization.

From this point of view, the blind user does not have to hold the exact external visualization in their heads [2]. Since the internal visualization (mental model) does not contain item specific information about the data, internal visualization is simulated by using carefully designed external data in the form of audio cues.

However, in certain cases, the internal visualization should be an exact replica of external visualization. For instance, in the image map, it is imperative to understand the position of each section. In addition, there are limitations in visualizing the parts as either stack or queue. Thus, to understand the external visualization or to expand the internal visualization, focus should be on interaction and not on the brain. In specific, the events performed by different parts of the body, such as hand and ear, which interact with external visualization to construct, manipulate and stimulate the mental model have to be understood.

IV. INTERACTION

The interaction technique is the technique used by the user to communicate with external visualization. It is characterized by 1) initiating physical action, and 2) following alterations in the visualization state. Tufte (1997) formulates around 57 interaction techniques to be used in the environment [15]. However, most of the interaction is ignored in HCI research since there are no appropriate features to be implemented with it. Thus, many features are aggregated to a single group.

Generally, interaction in touch screen technologies is classified into tactile and non-tactile interaction. Whereas clicking the button and holding the device is considered as tactile interaction, listening to audio is considered as non-tactile interaction. It can also be generalized based on characterization of visualization (static or dynamic) and based on modality (such as hand or ear interaction). HCI research reveals that human events are often ignored, which results in usability problems.

The main aspects of human events to consider regarding user interaction can be:

Project: Project is the special kind of action performed due to mental stimulation accompanied with keenness by shrinking the ear to concentrate. Krish (2009) observed that people give special attention to things which are augmented and projected [7]. The blind user is not able to project visual structure if the structure is visually projected by color encoding or circling. The project is a highly complicated task to project verbally through audio cues. It primarily depends on the intelligence of Text-to-Speech synthesizer and comprehension by the blind user. Although metrics are available to determine the intelligence and comprehension, context awareness is vital to understand the projection.

Select: The external representation is generated based on an algorithm in information visualization. The blind user creates a mental model of location while tutoring the system and subsequently modifies the model based on frequent

exposure and prior experience. The audio cues inform target name followed by the hand movements to reach the target.

Precision: An unexpected state of external visualization occurs if the interaction is not precise. Spatial factors such as cell padding, cell spacing and other related factors have an effect on the precision of hitting the target. The torque force and angle of inclination and recurrence of interaction also plays a vital role in hitting the target precisely.

Coupling: Occasionally, coupling of many actions gives rise to new events. It primarily occurs in multimodal interaction. For instance, a single click informs the target name through audio cues. On the other hand, if the blind user accidentally presses single click twice (equivalent to double click) this will invoke another event to open the item. Thus, the user has to get confirmation before proceeding to the next level to avert the unexpected event's happening.

Investigation: Investigation is a daunting task, especially for blind users if they don't have enough information at hand or any aural cue to reach the desired page. In this case, the blind user applies their own strategy in an ad hoc way to get the required result. When the exploration is new, the blind user uses an iteration process between discovered and new items. While in the investigation process, the blind user uses system help, domain expert or a sighted user and an existing related mental model if any, then the result will be the formulation of the new state of visualization. When the investigation does not achieve the required result, they alter the strategy.

Configuration: The blind user saves the state of visualization for later retrieval; whenever physical interaction takes place, the visualization is retrieved and applied. If a new feature of visualization is found, the new features are configured with the existing state to obtain an updated state of visualization. For instance, a user might identify a new way to reach a starting page.

V. PARTS OF BRAIN – A CASE STUDY

After a brief discussion about visualization, we present a case study to understand how an external visual space is understood by the blind user using a touch screen device. As part of our courseware, an image map is created for exploring the parts of the brain. Our techniques do not require alterations to the underlying touch screen hardware. These techniques are entirely software based. Our design aims to improve the accessibility of existing touch screen hardware for blind users.

We formulate two techniques, namely the Touch technique, and the Control technique. The touch technique is based on the Johnson-Laird model giving more emphasis the conceptual model, where the blind user has adapted to understand the system behaviour. In the control technique, more emphasis is given to user behaviour (Norman Model) in order to facilitate easy accessibility of the system. The techniques which we devised are discussed in detail now.

5.1 Touch based technique

According to these techniques a blind user has to press the surface of a touch screen either from left to right or vice versa or in a zig-zag manner. If the blind user touches the surface and the target is located, then the audio will inform the user of the name of the target (Fig. 3). If the target is not located, then no audio feedback will be received. Long tapping provides audio information about the target. Since we tested in small screen smart phone, if numbers of targets are more, then it is easy to locate more targets. If the target is less in the count and scattered wider around the space, then it is difficult to locate the target and time consuming. This technique needs a lot of patience and memorization of target location when used for the first time. Memorizing the target location will enable blind users to reach the target directly next time.

5.2 Control based technique

The control based technique is based on linearization of items while preserving the original layout of the screen. Blind users, on each tap over the widget placed on the bottom of the screen, will be provided with audio feedback about the name of the target (Fig. 4). A long tap will provide the audio information about the target as in the above techniques.

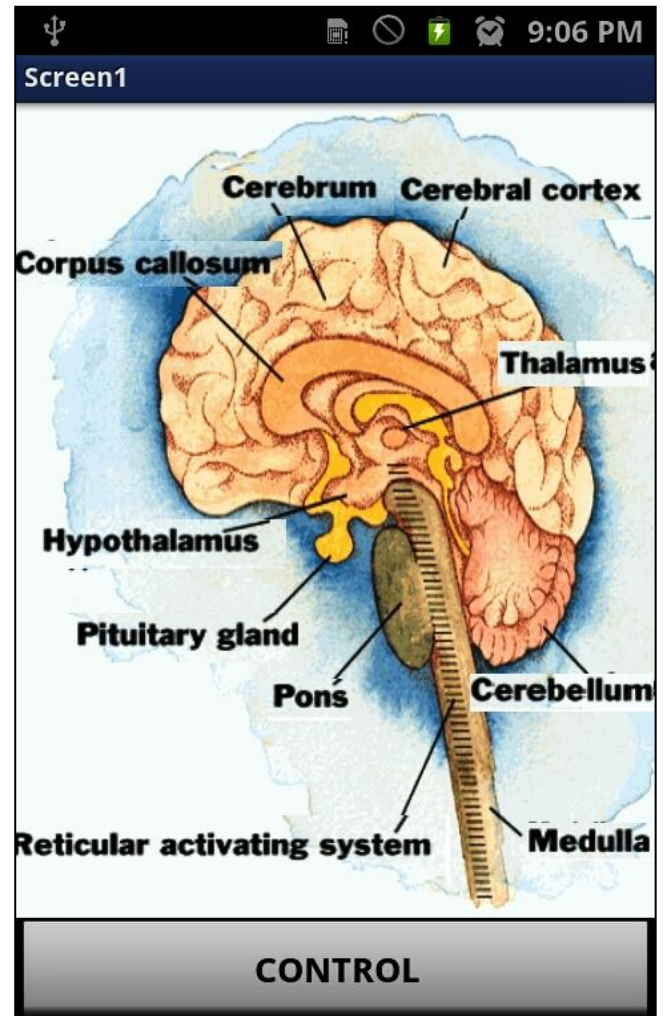


Fig. 4- Control technique

VI. EVALUATION

6.1 Participants

We recruited 10 blind computer users (3 male, 5 female), with an average age of 50.2 (SD=12.4). No participants used a touch screen before. However, all of the participants use a mobile using screen reader regularly in their daily life.

6.2 Apparatus

The study was conducted using Samsung Galaxy S2 smart phone based on Android. No hardware modification was made. We requested the blind users to identify the single widget above the home button in touch and control based techniques.

6.3 Procedure

The prototype is developed for each technique. The participants are tested to locate the item and understand the information about the item for each technique. Our prototype

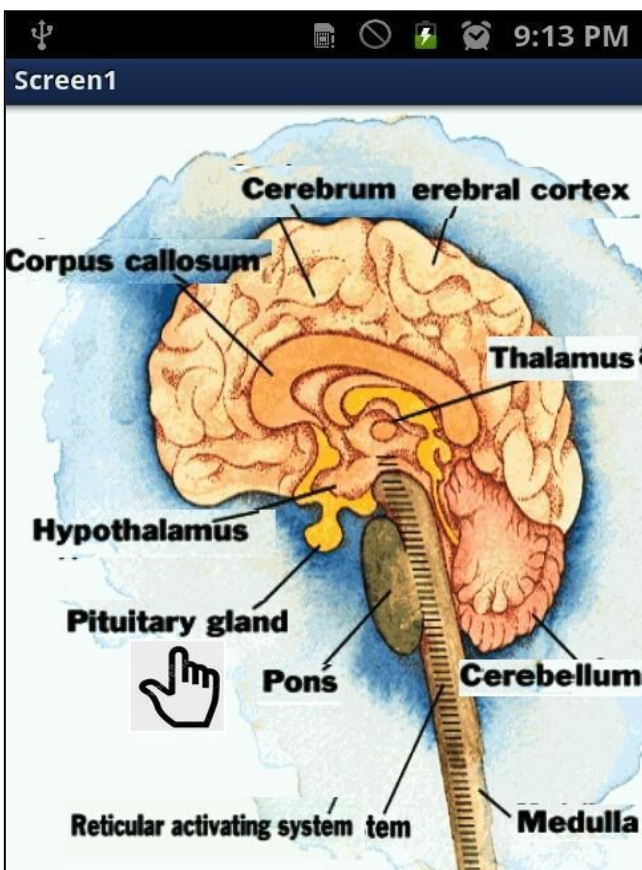


Fig. 3 – Touch technique

has an image map about parts of the brain. It has 9 items such as cerebrum, cerebellum, medulla, Pons and so on. The participants were given the target item and asked to reach the target item. While testing, we observe the following for each technique.

Task Acquisition Time: Time taken by the participant to reach the target was measured in seconds. The time starts when the participant performs the first tap and ends when he reaches the target item.

Stroke Count: Stroke count is the measure of the number of taps the participants performed to reach the target item.

After performing each technique, participants rated the technique using a Likert-scale questionnaire. After all techniques had been tested, participants ranked the techniques in order of preference. Questionnaires were administered verbally, and the experimenter recorded the answers.

VII. RESULTS

The user evaluation is conducted for the Quiz Touch prototype to evaluate accessibility and usability features.

Most complexity in dealing with touch screen by a blind user is in finding targets on the screen. A sighted user can quickly identify a target in a visual interface which uses empty spaces to design individual and group targets. Locating targets on the touch screen requires blind users to touch the empty area when they may not know where they are touching.

Each participant performed averagely 10 trials for each technique. The participant performance for each of the techniques and participant feedback is provided below. Our observation during the trials of each technique is also provided.

7.1 Descriptive statistics

We examined the target acquisition time and stroke count to reach the target for both the techniques.

Stroke Count: On average, blind participants used 21.17 strokes in touch technique to reach the target which is higher than the control technique for which they used 4.25 strokes on average to reach the target (Fig. 5). In addition, the maximum strokes performed by the blind user to reach the target were 60 in the touch technique and 9 in the control technique.

Task Acquisition Time: We analyze how much time the blind users required to reach the target. The descriptive statistics reveal that on average, 9.7 seconds are required to reach the target for the control technique. While the mean time for the touch technique is 43.91 seconds. For the control technique, the maximum duration to reach the target is 16 seconds and minimum is 2 seconds. The maximum task completion time in the touch technique is 125 seconds and minimum is 4 seconds.

7.2 Technique comparison

The dependent variables are checked for normality by using Shapiro Wilkson (W) test based on techniques. The data are not normalized for touch techniques using duration ($W(12) = 0.278, P < 0.05$) or control technique using stroke count ($W(24) = 0.191, P < 0.05$). Hence the Kruskal Wallis H test was performed for not normalized data to find the significance of each technique on the dependent variable.

From the result, it was concluded that there was a statistically significant difference between the duration of the techniques ($H(1) = 9.6134, P < 0.05$). It can be further concluded that the duration of touch (Mean rank=26.17) was more than the duration of control technique (Mean rank=14.67).

A statistically significant difference was also found between the techniques for a stroke count ($H(2) = 9.943, P < 0.05$). The stroke count for touch (Mean rank=26.21) was more than the stroke count of the control technique (Mean rank=14.65).

7.3 Feedback Analysis

The participants completed a questionnaire about the two techniques following the experiment. The participants show their compliance with each technique using a 5-point Likert scale (1=Strongly Disagree, 5=Strongly Agree) using a series of statements.

According to the Friedman Test, there was a statistically significant difference between the control and touch techniques, $\chi^2(1) = 45.302, p < 0.005$. Pair wise comparison using Wilcoxon on signed rank test found that the control technique is more preferred than the touch technique ($Z = -6.232, P < 0.05$).

Significant results were also found for the following measures: easy to use ($\chi^2(1) = 5, P < 0.05$), easy to learn ($\chi^2(1) = -2.60, P < 0.05$), familiar ($\chi^2(1) = -2.232, P < 0.05$), easy to navigate ($\chi^2(1) = -2.041, P < 0.05$) and intuitive ($\chi^2(1) = -2.06, P < 0.05$).

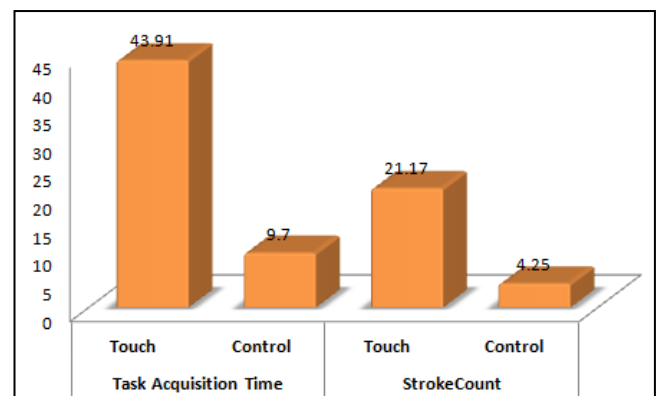


Fig. 5- Descriptive Statistics

VIII. DISCUSSION

A qualitative difference between the touch and control techniques with touch screen device is observed. The primary difference between the two techniques was how the blind user visualizes the external visualization. In the touch technique, users are required to scan the entire surface to locate the target. This method was somewhat slow and time consuming. Considerable efforts and patience are required to accomplish the task. Our observation of this technique reveals that blind users were frustrated during attempts to find the target.

On the other hand, in the control technique, blind users were able to navigate items in a linear fashion. This allowed the users to iterate the items quickly. In addition, the interaction with the system is minimized.

Although participants were faster overall with the control technique, they are not able to visualize the system. They are able to extract the external data through audio cues. However, the actual mind map between the data is missing. Using the control technique, the mind map is linear. On the other hand, using the touch technique, the blind user is able to visualize the external structure.

Note that we adopted “not” describing the actual location of each part. One important consideration when comparing the duration and stroke count is the feedback from the user. The feedback reveals that the control technique is more favoured irrespective of not achieving the visualization. In other words, the blind user prefers more behavioral aspects than conceptual aspects. Fortunately, they feel the visualization can be achieved through the domain expert (educators).

In the end, we can conclude that interaction plays a vital role in building internal visualization. The blind user expects easy interaction with the system. At the same time, the blind user prefers to build up their own mental model based on the external data. In order to maintain equilibrium with both the Norman and Johnson-Laird models, enough metadata has to be provided through external data either in the form of audio cues or through lecturing by educators. The metadata includes the actual location, size and shape of the external structure.

As a result, the external data are just converted to internal data and it is not of much utility in building the internal visualization equivalent to the external visualization. Specifically, the metadata about the external structure helps the blind users to build the internal visualization. By doing this, both behavioral and conceptual models can be kept in tandem.

IX. CONCLUSION

In this paper, we inspect the characteristics of internal representation relevant to an external system. We also investigate the interaction techniques useful to building internal visualizations when external visualization is based on the image map. Our research explores the solution for the questions mentioned in the first section.

What is meant by “internal visualization” for a blind user in terms of image map?. We distinguish the mental model as behavioural, structural and internal. In addition, the mental

model preservation data and mapping between the data or exact replicas of external visualization relies on metadata from external structures sent via audio cues.

For a given external visualization, how can it be related to internal visualization? The relationship between external and internal depends on the interaction technique. The interaction should be simple and easily accessible to the external system. The interaction may be based on the behavioural model developed by the user for easy access or a conceptual model developed by the designer to build the system and the blind user has to adapt to the system.

How do physical activities relate external visualization with internal visualization? It is necessary for external visualization to augment with internal visualization so that they form a blended system. Internal-external blending is performed in terms of six purposes: project, select, precision, coupling, investigation and configuration.

By addressing the research questions, we set up a framework to design an image map using touch screen that merges external visualization, internal visualization, interaction and analytical process. We trust this framework can direct and notify future actions on the design, evaluation and comprehension of the image map in courseware.

REFERENCES

- [1] Abidin, A. H. Z., Xie, H. & Wong, K. W., Blind users' mental model of web page using touch screen augmented with audio feedback. Kuala Lumpur, International Conference on Computer & Information Science (ICIS), pp.1046 – 1051,2012.
- [2] Brooks, R. A., Intelligence without reason. In: COMPUTERS AND THOUGHT, IJCAI-91. Morgan Kaufmann, pp.569-595, 1991.
- [3] Craik, K., The Nature of Explanation. Cambridge, Cambridge University Press, (1943).
- [4] Fakrudeen, M., Ali, M., Yousef, S. & Hussein, A. H., Analysing the Mental Model of Blind Users in Mobile Touch Screen Devices for Usability, pp. 837-842, (2013a).
- [5] Fakrudeen, M., Yousef, S., Hussein, A. H. & Ali, M., Presentation design of e-assessments for blind users using touch screen devices. Pitesti, pp. 1-6, (2013b).
- [6] Johnson-Laird, P. N., Mental models: towards a cognitive science of language, inference, and consciousness. Cambridge, MA, USA: Harvard University Press,1983.
- [7] Kirsh, D., Projection, problem space and anchoring. In Proceedings of the 31st Annual Conference of the Cognitive Science Society, pp. 2310–2315, 2009.
- [8] Liu, Z., Nersessian, N. & Stasko, J., Distributed Cognition

as a Theoretical Framework for Information Visualization. IEEE Transactions on Visualization and Computer Graphics, vol 14, pp.1173-1180, (2008).

[9] Murphy, E. et al., An empirical investigation into the difficulties experienced by visually impaired Internet users. Universal Access in the Information Society, Springer, Vol 7(1-2), pp.79-91, 2008.

[10] Norman, D., Some Observations on Mental Models. New Jersey: Lawrence Erlbaum Associates, 1983.

[11] Rouse, W. B. & Morris, N. M., On Looking Into the Black Box: Prospects and Limits in the Search for Mental Models. Psychological bulletin, VOL 100(3), pp. 349–363, 1986.

[12] Saei, S. N. S. M., Sulaiman, S. & Hasbullah, H., Mental model of blind users to assist designers in system development. IEEE CONFERENCE PUBLICATIONS, pp.1-5, (2010).

[13] Sasse, M., Eliciting and Describing Users' Models of Computer Systems, PhD, University of Birmingham, (1997).

[14] Scaife, M. & Rogers, Y., External cognition: how do graphical representations work?. International Journal of Human-Computer Studies, VOL. 45(2), pp. 185 – 213, Aug. 1996

[15] Tufte, E. R., Visual Explanations: Images and Quantities, Evidence and Narrative. 4th printing with revisions edition ed, Graphics Press, 1997.

[16] Zhang, Y., The influence of mental models on undergraduate students' searching behavior on the Web. Information Processing & Management, VOL. 44, pp.1330–1345, 2008

An Implementation of Android-based Steganography Application

Sultan Zavrak

Dept. of Computer Engineering
Duzce University
Duzce, Turkey

Seyhmus Yilmaz

Dept. of Computer Engineering
Duzce University
Duzce, Turkey

Huseyin Bodur

Dept. of Computer Engineering
Duzce University
Duzce, Turkey

Abstract— The rapid development of smart phone technology has led to cheapening of the phone equipped with many advanced features such as sensors. One of the most widely used sensor in the phone is its camera. Although the photographs captured by camera can be shared via many ways, one of the most commonly used sharing methods is Multimedia Message Service (MMS) which allows transmission of files such as photographs, audio and video. A major disadvantage of MMS is that it doesn't provide sufficient safety mechanism and because of this, the data of the people who wants to hide confidential information from state-controlled systems can be easily monitored. In this study, unlike cryptography-based, a steganography-based mobile application that can embed the confidential information into an image, send it to receiver, and extract the confidential information from the image in the receiver side is developed. Besides, the performance data such as the embedding and extraction time of confidential information and experimental results of application are given.

Keywords— Multimedia Message Service, MMS, Steganography, Smart Phone.

I. INTRODUCTION

In many parts of the world, the rapid growth of smart phone technology makes smart phones equipped with many advanced features such as accelerometer, gyro, proximity, and compass sensors cheaper. One of the most widely used of these features is the phone's camera. Smartphone cameras have reached a quality that can compete with professional cameras.

The images captured by the camera are often shared by users on social networking websites such as Facebook, Instagram etc. As an alternative to those websites, and also older technology, users utilise Multimedia Message Service, which does not require the use of the internet. MMS also provides users to share media files such as video and audio. The most popular uses are to send photos from mobile devices equipped with cameras.

A various media format is supported by MMS. No less than two codecs can be compatible such as aacPlus and AMR-WB. Compatibility for MP3, MIDI and WAV file extensions can be recommended as well. In addition, MMS customers ought to support JPEG with JFIF for still pictures, whilst the compatible extensions are GIF87a, GIF89a and PNG for bitmap graphics.

Lastly, it supports media extensions H.263 Profile 3 Level 45, MPEG-4 Visual Simple Profile Level Ob and H.264 (AVC) Baseline Profile Level lb. Furthermore, the main disadvantage is that the size of media that should be less than 30kB, a picture should not be more than 100kB and a film has to be less than 300kB [1, 2].

A standard method to keep a confidential message is that the cryptography which is the method to encode information into completely varied type. The confidential information may be retrieved solely by recognizing the way to come them to the first type. Another method to keep confidential information is steganography which is a technique of concealing the confidential message inside of another data such as text, audio, image or video. Cryptography is that the study of concealing info, whereas Steganography deals with composing hidden messages so solely the sender and therefore the receiver recognize that the message even exists. In Steganography, solely the sender and therefore the receiver recognize the existence of the message, whereas in cryptography the existence of the encrypted message is visible to the planet. Because of this, Steganography removes the unwanted attention returning to the hidden message. Cryptographic ways attempt to defend the content of a message, whereas Steganography uses ways that may hide each the message moreover because the content. The limitations of the encryption are identical for Steganography as well, mutually they need that sender and recipient formerly agree on the accepted technique. A confidential message was formerly swapped through a safe communication between the sender and the recipient [1].

MMS is a convenient service for hiding secret message with stenographic methods as it provides image, video and audio transmission as well as text transmission. Secret message can be sent to other people by embedding this message into an image, a video or an audio file.

In this study, images have been chosen as an embedding medium because photo sharing is widely used by people while text has been used as a secret message. The disadvantage of the MMS is that the size of messages is limited like the same as in SMS. Therefore, images captured by smart phone's camera should not exceed 100 KB to be embedded confidential data into them.

The aim of this work is to avoid the computational power of smart phones and other hardware features of the problems

raised by the previously [1] and [2] studies with the development of smart phone technology.

The rest of the study is organized in this way. In the first part, we illustrate F5 steganography method that we used in the application by its giving algorithmic steps. The subsequent section explains the mobile application that builds stenographic approach on a smartphone, and we evaluate its performance. In the end, we give the concluding remarks.

II. F5 STEGANOGRAPHY

In this paper, the implementation of steganography to images which have taken from a camera of Smartphone has been presented. Because of the amount of the resulting files, the raw bitmap form of pictures is not generally utilized on smart phones. As an alternative, the JPEG form developed into an extensively adopted compression form to keep and transmit photos [8]. Embedding confidential information into a JPEG format is very difficult. This is because such file is lossy compressed type, and it is not entirely the identical photos.

Thus, the approach presented by the JPEG compression may demolish confidential information straight inserted into the raw bitmap. Additionally, changing a JPEG picture with no extraordinarily altering the coded picture nearly is not possible. JPEG pictures have a predefined framework because of the compression algorithm. As a result, a small alteration of the picture may also be detected by presenting the encrypted picture.

The general diagram of steganography process illustrated as in Fig. 1.

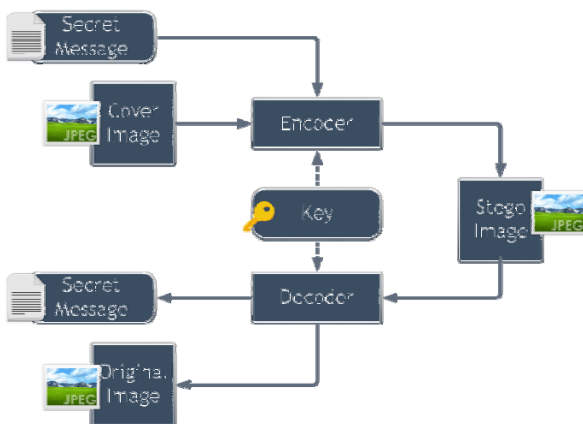


Fig. 1 Steganography process diagram

Although most stenographic methods provide a high capability for confidential information, they are vulnerable to visual and statistical attacks. Implementations that are resistance to such attacks just provide a very little capability.

Confidential information is embedded in lossy compressed JPEG photos by Jsteg, which is a stenographic kit. The capability of this kit is high. In addition, such kit is resistance to visual attacks. Nevertheless, alterations made by Jsteg can be found out by statistical attacks [3]. Acoustical and visual attacks can be defended by MP3Stego and IVS Stego correspondingly [4]. Outside of this, the greatly small

embedding ratio stops each recognized statistical attack. The F4 method endures against both visual and statistical attacks by combining both preferences as well as high capacity. In the most cases, the embedded confidential information is less than the cover media capacity. For that reason, part of it is not changed.

Steganalysis attackers can exploit such discontinuity. In order to prevent hinder such attacks, the confidential information is scattered by the F5 algorithms over the entire carrier medium. At first the coefficients are shuffled by the F5 using a permutation. After, the confidential information is embedded in the permuted series. The permutation is reliant upon a key that is generated from a password. After, the changed coefficients are sent into their real sequence, to the subsequent stage of the JPEG coding by the F5. In order to calculate the equal permutation of the coefficients and see the confidential information, the recipient of the enciphered information requires the password [5].

The F5 algorithm follows the coarse structure below [4].

Step 1. Initiate compressing the JPEG and then halt the quantization of coefficients.

Step 2. A cryptographically robust random numeral producer with the key generated by the code is started.

Step 3. Two factors, random producer and amount of coefficients, are produced.

Step 4. Define “ k ” from the carrier medium capability, and the size of the confidential information.

Step 5. Compute the size of the password “ $n = 2^k - 1$ ”.

Step 6. The confidential information is embedded with $(1, n, k)$ matrix code.

a) Filling the buffer by using nonzero coefficients.

b) Create a hash result number with k bit length.

c) Insert the subsequent k bits of the secret to the hash result (bit by bit, XOR).

d) If the total is zero, the buffer remains the same. If not the total is the order of the buffer $1, \dots, n$. The exact result of its factors must be decreased.

e) Check for shrinkage. If there is shrinkage, alter the buffer (remove zero reading 1 extra non-zero coefficient, i. e. do again Step 6a starting from the identical coefficient). When no shrinkage, forward to fresh coefficients after the real buffer.

When confidential message is carried on with Step 6a.

Step 7. Carry on compressing JPEG.

III. IMPLEMENTATION

We designed and implemented the application for Android-based smartphones because it is widely used mobile operating system in the world [6]. The application consists of three interfaces. The first one is the main interface, which is shown in Fig 2, which directs the user the embedding or extracting interfaces.

The embedding interface, as illustrated in Fig. 3, consists of several components to get necessary parameters such as

secret text, password and the image to be used for embedding from the user. Hidden text is embedded into an image that the user chooses. The selected pictures can be taken from the currently selected camera or one of the existing local storage devices. Before embedding hidden text into the pictures, the user should check that the file size for images does not exceed 100KB. If the file size exceeds 100KB it should be scaled to the appropriate size before continuing to proceed.

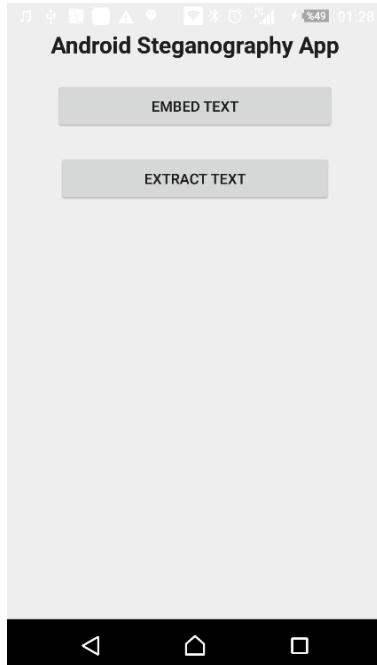


Fig. 2 The main interface of the application

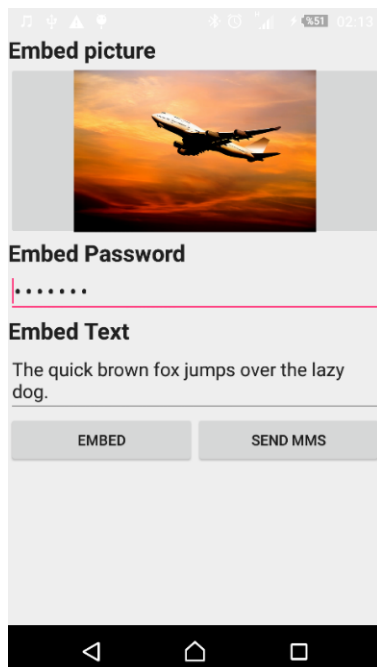


Fig. 3 Embedding interface of application

The processing steps performed are shown below.

1. The text that will be embedded into the picture should be entered.
2. The password that will be used should be entered.
3. A picture is captured by smart phone and a file is taken from the memory. The file size should be less than 100 KB. If it exceeds that size, it will be resized.
4. The secret text is embedded into picture by using F5 method.
5. The picture is sent to the recipient as MMS.

The extracting interface, as illustrated in Fig. 4, consists of same components to get necessary parameters except secret text from the user.

IV. PERFORMANCE EVALUATION

We tested our android implementation on a smart phone Sony Xperia Z3, running under the Android 5.1.1. The representative of existed the smartphones on the shelves is the Z3. This provides computation power compared to currently existing smart phones. In general, the structure used by smart phones relied upon CPU ARM9 at Quad-core 2.5 GHz, and the RAM used by the Z3 is approximately 3 GB. Although the screen of the Z3 uses a widespread size of 1920x1080 pixels their performance is quite low in compared to other ones duo to the 16 million appearing colours.

A common 20.7 MegaPixels, 3840 × 2160 pixels for the built-in photo camera, which can provide 24 bit colours, is used.

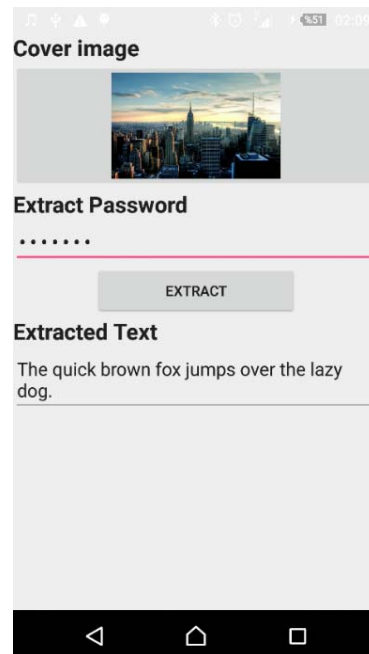


Fig. 4 Extracting interface of application

The selected pictures were intentionally taken from different environments. In this way, images with different colour levels have been tested. We have four different groups:

sea, forest, sea and city are shown in Fig. 4. Also, some features such as size, width, height are given in Table 1.



Fig. 5 The sample photos

TABLE I
THE PROPERTIES OF TEST IMAGES

| Picture Name | Width (Pixels) | Height (Pixels) | Bit Depth | Size (KB) |
|--------------|----------------|-----------------|-----------|-----------|
| Airplane | 700 | 467 | 24 | 94 |
| Forest | 500 | 281 | 24 | 81,5 |
| Sea | 700 | 332 | 24 | 87,3 |
| City | 500 | 313 | 24 | 71,1 |

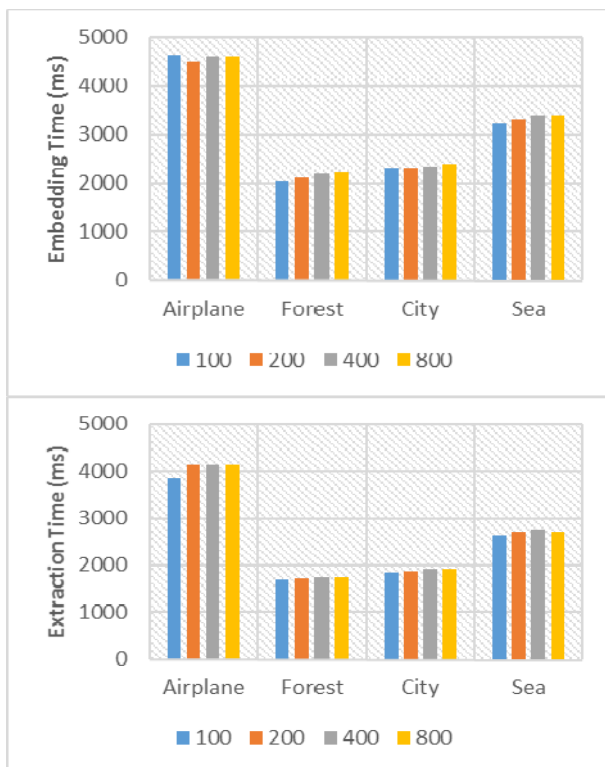


Fig. 6 Time graphs of embedding and extraction operations

In our tests we measured the time to embed and extract a secret message into a cover media. We used four data, whose size was 100, 200, 400, and 800 number of characters. We tested the program on several cover photos, stored as JPEG files; some of them are shown in Fig. 5.

As you can understand from the graphs in Fig 6, both embedding and extraction time of confidential information are proportional to width and height of the images. At the same time, we can reach a conclusion that the character count of secret text doesn't affect times of both operations very much.

V. CONCLUSION

In this study, we show that the steganography can be used in smartphones to hide confidential data from other unauthorized users.

This method is not restricted to smart phones, since MMS messages also can be sent through electronic mail or other instant messaging applications such WhatsApp. Consequently, this technique can be used on other devices such as PCs and tablets.

REFERENCES

- [1] A. Amoroso and M. Masotti, "Lightweight steganography on smartphones," *Consum. Commun. Netw. Conf.*, pp. 1158–1162, 2006.
- [2] M. Shirali-Shahreza, "Steganography in MMS," *INMIC2007 - 11th IEEE Int. Multitopic Conf.*, pp. 5–8, 2007.
- [3] A. Westfeld and A. Pfitzmann, "Attacks on Steganographic Systems," in *Proceedings of the Third International Workshop on Information Hiding, 2000*, pp. 61–76.
- [4] A. Westfeld, "F5—A Steganographic Algorithm," in *Information Hiding SE - 21*, vol. 2137, I. Moskowitz, Ed. Springer Berlin Heidelberg, 2001, pp. 289–302.
- [5] A. Westfeld, "F5 — A Steganographic Algorithm High Capacity Despite Better Steganalysis," in *Lecture Notes in Computer Science, 2001*, pp. 289–302.
- [6] "IDC: Smartphone OS Market Share 2015, 2014, 2013, and 2012." [Online]. Available: <http://www.idc.com/prodserv/smartphone-os-market-share.jsp>. [Accessed: 20-Dec-2015].

Survey of the Adaptive QoS-aware discovery approaches for SOA

Monika Sikri #¹

Cisco Systems India Pvt. Ltd.

SEZ, Embassy Tech Village, Panathur Devarabeesanahalli, Bangalore East Taluk Bangalore India 560037

¹msikri@gmail.com

Abstract—Service Oriented Architecture is very commonly used as an architectural paradigm to model distributed integration needs. It provide the means of achieving organizational agility by building applications, which adapt to the dynamic needs of the business. Agility to Adapt is one of the key drivers for its growth. Its widespread adoption has led to the proliferation of multiple services offering similar functionality but different Quality of Service on the enterprise network. In real-time enterprise environment services are added and removed from the network on the fly. The service discovery approach does not only need to consider the QoS of other similar services but also the impact of dynamic and unpredictable system behavior on QoS. In view of this, there is a need for adaptive discovery approach that can accommodate these run-time changes and keep the system functioning despite the unpredictability. As part of this work we have reviewed existing works in Adaptive QoS-aware discovery for SOA systems to understand the gaps and need for future research.

I. INTRODUCTION

Followed by its huge adoption at an application level the scope of service-orientation has encompassed business models offering capabilities as services. SOA acts as a driver for agile adaptation, required to support the organic and inorganic growth [1] led by mergers and acquisitions, change in business scope or marketplace dynamics. With this widespread adoption and growth of SOA it is common to see multiple services offering similar functionality but different QoS. Discovery of the unique service, out of multiple functionally similar services requires a framework and QoS model to enable QoS-aware selection and discovery to support the same.

During its formalization the key benefits of SOA were mainly alluded to interoperability and loose-coupling. Traditional SOA could sustain the monolithic development methodology, static QoS-aware systems. However in view of this widespread adoption there is a need for renewed focus on Quality of Service (QoS) aspects. The existing SOA deployments can support some of the dynamic discovery needs for these redundant service like location transparency enabled by shared enterprise registries like UDDI, rule-based recovery, dynamic monitoring based update of QoS attributes using mediation support as provided by Enterprise Service Bus(ESB) [2]. These kind of approaches can support predictable recovery actions and responses from the system.

There is a need for current enterprise SOA deployments to scale adaptive to enable the next phase of Adaptive Service Oriented Enterprise (ASOE). For a distributed paradigm as SOA, adaptation cannot always be externalized from the system. It needs to be engrained in the architecture to enable autonomic behavior corresponding to run-time and unpredictable changes. As part of this work we have specifically reviewed the works on Adaptive QoS-aware discovery for SOA systems.

It is realized that Adaptive QoS-aware discovery is being modeled as an optimization problem in the literature. An optimization problem is the one that has more than one feasible solution. It uses an objective function to choose a unique solution for the given set of constraints. It can be mathematically represented as follows:

The standard form of an optimization problem is:

$$\underset{x}{\text{minimize}} \quad fO(x) \quad (1)$$

$$\text{subject to} \quad a_i(x) \leq 0, \quad i = 1, \dots, m \quad (2)$$

$$b_i(x) = 0, \quad i = 1, \dots, p \quad (3)$$

Equation 1 shows the objective function that need to be minimized. Equation 2 shows the inequality constraints. Equation 3 indicates the equality constraints. The problem need to be resolved in a way to minimize the value of the objective function under the given set of constraints. The solution to an optimization problems offers two possible minimas for the objective function known as: local and global. Local minima is a point where its value is smaller than or equal to the value at neighborhood. On the other hand in case of global minima, this minimum value is less than all other feasible minimum value for this function. Modeling this problem for QoS-aware discovery can offer two possible approaches for service selection: local and global. The local selection of the service results in finding the service satisfying all the local constraints within the class. The global selection finds the service that meets the constraints end-to-end in the whole problem domain. The problem domain refers to the scope of the problem. The next section explains the state of the art discovery approaches in Adaptive QoS-aware discovery in SOA systems.

II. STATE OF THE ART

This section discusses the works in adaptive QoS-aware web service discovery to understand the current state [3] and the need to address further research. As discussed QoS-aware discovery is being solved to provide two kinds of optimization approaches: global or local. The global optimization approach for QoS-aware discovery can get very complex for dynamic environments due to run-time changes in QoS. The service composition and business process layer for SOA services requires dealing with composite workflows and their integrations. For a given workflow selecting all the services satisfying the global QoS constraints requires an exhaustive computation. This problem can be modeled as an instance of combinatorial problem known as multiple-choice knapsack problem (MMKP), which is a strong NP-hard problem. The global optimization approach need to be developed in a way to restrict the exhaustive listing and computation. The solution to this problem requires modeling the optimization in a way to apply restrictions and to limit and prune the number of options to reduce the problem domain. Many authors have used mixed integer programming and constraint programming approaches for service selection to solve these problems.

Realizing the unprecedented opportunity for B2B collaboration through web service during the conceptualization of Web Service standards WSDL and SOAP, an adaptation need was realized to deal with dynamic run-time changes for quality aware SOA systems. Zeng et al. [4] proposed AgFlow as a middleware platform for QoS-aware web service composition. The platform suggested a multi-dimensional QoS model, global and local service selection approaches and an adaptive engine for changing the execution by re-planning during the execution of composite service. The architecture in AgFlow has three main components: a service broker, a service composition manager and web services. The service brokers allows registration of service providers information in the UDDI registry. The service providers information include QoS and location of the service. The service composition manager component contains an execution engine. It contacts the service broker to fetch Web Service location information. It uses this information to select and execute service composition. For web service composition the statecharts [5] are used to represent dependencies in web service composition. The simple additive weighting is used by merging the quality vectors to build a quality matrix representing the best service for local optimization. The global optimization approach used mixed integer programming to determine the selection based on the frequency of the path used to prune the possible number of paths. The statecharts used in this work by default models reactive systems which respond to events. The adaptiveness is achieved by choosing the best service using the weighted approach limited by QoS values.

A combined optimization and negotiation approach considering the statefulness of the web service is proposed in [6] for adaptive service composition for large flexible business processes with severe constraints. This approach is

implemented using MAIS (Multichannel Adaptive Information Systems) [7] architecture, that allows execution of flexible business processes. This architecture allows selection of services for flexible business processes at run-time. The concrete information about services is obtained from MAIS registry that provides the service interface and QoS information. The selection is being modeled as an optimization problem using mixed integer linear programming (MILP) problem by linearizing the constraints and objective function. If the feasible solution is not found using the optimization approach negotiation mechanism is suggested between service brokers and providers to pump service provider to provide better QoS if needed. It increases the solution domain of the problem by introducing more services satisfying the negotiation and optimization criteria. As an added functionality authors have provided a new set of global constraints for considering the state of the services. Though it considered the severity of constraints for large flexible business process the solution is difficult to scale for large traffic on the single service. Also the changes in service behavior due to changing quality profile needs to be investigated under varying load conditions.

A heuristics based service selection approach has been suggested in [8] for service composition satisfying global constraints. The approach uses a broker based architecture (QBroker) to enable selection of the best service. QBroker acts as a intermediary to perform service discovery, planning, selection, and adaptation. It follows a sequence of stages for service selection and composition. It first creates a process plan that defines an abstract process flow of component functions and their relationships followed by a functional graph. All possible process plans fulfilling a user request are represented in the function graph. Finally a service candidate graph is built from functional graph. It includes path for all combinations of function nodes that is used to fulfill the service request. The two heuristics based approaches have been suggested based on the type of workflow termed as WS-HEU and WFlow. WS-HEU approach is suggested for sequential workflow. WFlow approach is used for generic workflows. Multidimensional multichoice 0-1 knapsack problem (MMKP) represents the combinatorial model of the problem and Multi-constraint optimal path (MCOP) represents the graph view of the problem. Both WS-HEU and WFlow are being modeled using combinatorial and graph based model. It uses Stochastic workflow Reduction (SWR) algorithm to shrink the workflow until one task is left. Though this solution provides the near optimal solution it will not scale to the need of business critical services requiring real-time on-the-fly handling of QoS changes.

Processes with Adaptive Web Services (PAWS) [9] offers a framework for flexible and adaptive service-based processes design-time and run-time adaptation flexible service-based. It exhibits self-optimization and self-adaptation in the business process execution. The authors have emphasized capturing the adaptation characteristics at design-time to capture QoS constraints and annotated process definition to retrieve services satisfying functional and non-functional requirements. The

authors have realized that automating negotiation at design time saves time during run-time for process optimization. The process definition at this point of time is abstract. The run-time modules performs process optimization while selecting the services based on annotated requirements of local and global constraints. The service retrieval takes place through URBE [10], which is proved to be 80 percent accurate in determining similarity. A mixed-integer linear programming (MIP) approach is used for service selection using global QoS process constraints, and the runtime execution context.

Mabrouk et al. [11] have proposed a solution for adaptive QoS-aware composition using heuristics algorithm for service discovery. The authors have used semantic-based discovery approach suggested in the work [12],[13] for service discovery. It uses domain-specific language and ontologies to match the candidate services for a given composition. As a first step in service discovery filtering is done based on advertised QoS declared by providers. As a next step it determines the service composition based on using user-defined global constraints adopting a heuristics based approach. That is, it selects a set of service candidates for each abstract activity. The service selection is done using K-means [14] heuristics based approach [15] that, when composed together, meet global QoS constraints.

A hybrid solution for service composition has been proposed in [16] for end-to-end QoS constraints. This solutions combines global optimization approach with local selection by mapping global constraints into local constraints. Once the global constraints are obtained a distributed search is performed to obtain the required service. The qualified sets of services are sorted based on the utility function. The QoS composition problem is represented with the four constructs : sequential, loop, parallel, conditional. A generic QoS model representation is used for domain specific and independent attributes using quantified value is assumed. First a QoS vector of the composite service is defined followed by the aggregation functions for the quality attributes. It is seen that quantifies QoS attributes can be represented as representing summation, product and minimum and maximum value representation. A utility function to determine the ranking based on that. The two phases of ranking involves scaling and computation. At first the multi-dimensionality of QoS attributes is resolved using the scaling mechanism followed by a weighting process to represent priorities and user preferences while evaluating the ranking of candidate services at first the The minimum and maximum QoS value is obtained by aggregation of these values from each service class. These values are then used in utility function for computation. For each service class local quality levels are computed to use them while mapping global constraints to local. Unlike other works authors have modeled the constraints decomposition as MIP while the search as linear problem within the subclass.

A mathematically sound framework for QoS-aware management of service orchestration has been proposed in [17] with special focus on data-dependent workflows. The other works have assumed and addressed this issue in computation

and proposed approach. However in this work the authors have specifically proposed a mathematically sound framework for contract-based management of service orchestrations using probabilistic QoS. The data dependent workflows and changing QoS result in non-monotonic orchestrations with respect to QoS. This means that improving QoS of an individual service may actually reduce the composite service aggregated QoS. Due to this there will be a difference in local vs global optimization as being also [18], [19]. The QoS composition rules are defined using mathematical calculus defined as part of this work. The authors have used probabilistic QoS to deal with uncertainty. In the literature Markov models [20], [21] are generally used to deal with probabilistic QoS. However due to uncertainty and non-determinism in data dependent workflows the authors have used Markov Decision Model [22] to consider rest of the decision. In the nutshell, the problem is being modeled as Stochastic Control Problem [23], which optimizes the remaining overall QoS in the workflow. The authors have rightly realized the importance of monotonicity and provided the mathematical framework in a way to provide monotonicity for these orchestrations.

Orchnet is developed as an extension to Petri net based modeling [24] as a model for QoS-aware service composition. The work done by authors does formalize and establish the importance of monotonicity in service orchestrations mathematically enabling similarity in resolving the problems due to ease of equivalence in local and global optimization for monotonic orchestrations.

The Table I summarizes the works discussing adaptive service discovery based on non-functional properties in SOA systems.

For the adaptation the change in QoS values is obtained based on the feedback from the system. There are various ways to provide feedback [25] like fuzzy logic [26], prediction-based [27], reputation-based [28], [29]. Once the feedback is obtained, an approach for ranking the service is required using the same. This feedback for QoS values could be an exact score as in reputation-based mechanism or need further computation to determine the rating. The approach of quality vector aggregation has also been proposed in some like [4] and [16]. In [4] QoS data retrieval mechanism differs based on the QoS dimension of the attributes. The work in [16] calculates aggregated quality vectors for composite tasks in services based on their relationship in the workflow. The next section explains and concludes the work identifying future scope.

III. CONCLUSION

This paper specifically focuses on understanding the current state of Adaptive QoS-aware discovery in SOA systems for redundant service deployments. It becomes inadvertent to model service discovery architecture in scenarios where multiple redundant services are available on the enterprise network serving similar functional but different non-functional characteristics represented by QoS attributes. It is observed that the discovery approach gets varied based on the scope of

| Work | QoS specification | Intent | Optimization |
|------|---|--|------------------|
| [4] | An extensible QoS model is defined for multi-dimensional QoS. It has been represented for mainly five QoS criteria. | Ag-Flow: A middleware platform that performs optimization and re-configuration of service composition. | Local and Global |
| [11] | A semantic QoS model formulated as a set of ontologies including QoS | A near optimal solution for QoS-aware service composition in dynamic environments | Local and Global |
| [9] | A generic QoS model is specified to define QoS characteristics. | PAWS: A framework for business process optimization and adaptation for changing contexts | Global |
| [6] | An extensible QoS model is defined based on the aggregation pattern and negotiability | Adaptive service composition and optimization | Global |
| [8] | A generic QoS model with functions defined for service composition | Service selection for composite services with end-to-end QoS constraints | Hybrid |
| [16] | A generic QoS model has been assumed with discrete QoS values | Dynamic service composition based on client constraints | Hybrid |
| [17] | A probabilistic model for multi-dimensional QoS is used to handle monotonicity | QoS-aware management of modeling monotonic service orchestrations for data dependent workflows | Local and Global |

TABLE I: Literature Survey for Adaptive QoS-aware approaches for SOA systems.

the problem. This scope in turn depends on the operational execution layer in the SOA stack. It means it is addressed differently for basic, composite or business process layer. The discovery problem is solved as an optimization problem providing local and global solution for service selection. The local approaches for discovery are effective but do not offer the best solution to satisfy end-to-end global constraints for example maximum availability, minimum price. Recently there has been focus on the global optimization approaches. In fact the second generation solutions [9] are mostly offering global optimization approaches. It is observed that most of these works focus on the reconfiguration, optimization and adaptation in the business process orchestration and service-composition layers. In these cases there is an assumption of a pre-defined workflow. For a given workflow the local optimization approach selects the best service for a given task in the neighborhood satisfying the local constraints whereas the global optimization approach explores all possible combinations to satisfy global end-to-end constraints. In real-world, the number of similar services in these cases may not be very large and hence solutions can be adapted to scale. However we find an exclusive need to address the issue for foundational services layer or platform services where replication and redundancy is one of the key ways to improve QoS. A solution designed for these fundamental services can scale

ACKNOWLEDGMENT

I would like to thank my organization Cisco Systems Inc., for the continuous support for our research. I would also like to thank BITS Pilani for their guidance and reviews on research directions.

REFERENCES

- [1] K. S. Boppana, "Enterprise software: analysis of product strategies," Ph.D. dissertation, Massachusetts Institute of Technology, 2006.
- [2] D. A. Chappell, *Enterprise Service Bus: Theory in Practice*. Sebastopol, CA: O'Reilly Media, 2004.
- [3] J. Choi, D. Nazareth, and H. Jain, "The impact of SOA implementation strategies on business value and agility: A systems dynamics approach," *2010 6th International Conference on Advanced Information Management and Service (IMS)*, pp. 1–6, 2010.
- [4] B. Benatallah, A. Ngu, M. Dumas, J. Kalagnanam, and H. Chang, "QoS-aware middleware for Web services composition," *IEEE Transactions on Software Engineering*, vol. 30, no. 5, pp. 311–327, may 2004. [Online]. Available: <http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=1291834>
- [5] D. Harel and A. Naamad, "The STATEMATE semantics of statecharts," *ACM Transactions on Software Engineering and Methodology*, vol. 5, no. 4, pp. 293–333, oct 1996. [Online]. Available: <http://dl.acm.org/citation.cfm?id=235321.235322>
- [6] D. Ardagna and B. Pernici, "Adaptive Service Composition in Flexible Processes," *IEEE Transactions on Software Engineering*, vol. 33, no. 6, pp. 369–384, jun 2007. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1263152.1263531>
- [7] P. Atzeni, T. Catarci, and B. Pernici, "Multi-channel Adaptive Information Systems," *World Wide Web*, vol. 10, no. 4, pp. 345–347, jun 2007. [Online]. Available: <http://link.springer.com/10.1007/s11280-007-0034-9>
- [8] T. Yu, Y. Zhang, and K.-J. Lin, "Efficient algorithms for web services selection with end-to-end qos constraints," *ACM Trans. Web*, vol. 1, no. 1, May 2007. [Online]. Available: <http://doi.acm.org/10.1145/1232722.1232728>
- [9] D. Ardagna, M. Comuzzi, E. Mussi, B. Pernici, and P. Plebani, "PAWS: A framework for executing adaptive web-service processes," *IEEE Software*, vol. 24, no. 6, pp. 39–46, nov 2007. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4375241>
- [10] P. Plebani and B. Pernici, "URBE: Web service retrieval based on similarity evaluation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 11, pp. 1629–1642, nov 2009. [Online]. Available: <http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=4760142>
- [11] N. B. Mabrouk, S. Beauche, E. Kuznetsova, N. Georgantas, and V. Issarny, "Qos-aware service composition in dynamic service oriented environments," in *Proceedings of the 10th ACM/IFIP/USENIX*

- International Conference on Middleware*, ser. Middleware '09. New York, NY, USA: Springer-Verlag New York, Inc., 2009, pp. 7:1–7:20. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1656980.1656990>
- [12] S. Ben Mokhtar, A. Kaul, N. Georgantas, and V. Issarny, "Efficient semantic service discovery in pervasive computing environments," in *Middleware 2006*, ser. Lecture Notes in Computer Science, M. van Steen and M. Henning, Eds. Springer Berlin Heidelberg, 2006, vol. 4290, pp. 240–259. [Online]. Available: http://dx.doi.org/10.1007/11925071_13
- [13] S. B. Mokhtar, D. Preuveneers, N. Georgantas, V. Issarny, and Y. Berbers, "Easy: Efficient semantic service discovery in pervasive computing environments with qos and context support," *J. Syst. Softw.*, vol. 81, no. 5, pp. 785–808, May 2008. [Online]. Available: <http://dx.doi.org/10.1016/j.jss.2007.07.030>
- [14] S. Z. Selim and M. A. Ismail, "K-means-type algorithms: a generalized convergence theorem and characterization of local optimality," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 1, pp. 81–87, 1984.
- [15] S. P. Lloyd, "Least squares quantization in pcm," *Information Theory, IEEE Transactions on*, vol. 28, no. 2, pp. 129–137, 1982.
- [16] M. Alrifai, T. Risse, and W. Nejdl, "A hybrid approach for efficient Web service composition with end-to-end QoS constraints," *ACM Transactions on the Web*, vol. 6, no. 2, pp. 1–31, may 2012. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2180861.2180864>
- [17] A. Benveniste, C. Jard, A. Kattepur, S. Rosario, and J. A. Thywissen, "QoS-aware management of monotonic service orchestrations," *Formal Methods in System Design*, vol. 44, no. 1, pp. 1–43, jul 2013. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2581928.2581968>
- [18] M. Alrifai and T. Risse, "Combining global optimization with local selection for efficient qos-aware service composition," in *Proceedings of the 18th international conference on World wide web*. ACM, 2009, pp. 881–890.
- [19] D. Ardagna and B. Pernici, "Global and local qos guarantee in web service selection," in *Business Process Management Workshops*. Springer, 2006, pp. 32–46.
- [20] L. Abeni and G. Buttazzo, "Qos guarantee using probabilistic deadlines," in *Real-Time Systems, 1999. Proceedings of the 11th Euromicro Conference on*, 1999, pp. 242–249.
- [21] S. Gallotti, C. Ghezzi, R. Mirandola, and G. Tamburrelli, "Quality prediction of service compositions through probabilistic model checking," in *Quality of Software Architectures. Models and Architectures*, ser. Lecture Notes in Computer Science, S. Becker, F. Plasil, and R. Reussner, Eds. Springer Berlin Heidelberg, 2008, vol. 5281, pp. 119–134. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-87879-7_8
- [22] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.
- [23] H. J. Kushner, *Introduction to stochastic control*. Holt, Rinehart and Winston New York, 1971.
- [24] P. Xiong, Y. Fan, and M. Zhou, "A Petri Net Approach to Analysis and Composition of Web Services," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 40, no. 2, pp. 376–387, mar 2010. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5398972>
- [25] J. M. García, M. Junghans, D. Ruiz, S. Agarwal, and A. Ruiz-Cortés, "Integrating semantic Web services ranking mechanisms using a common preference model," *Knowledge-Based Systems*, vol. 49, pp. 22–36, sep 2013. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2501252.2501300>
- [26] V. X. Tran, "WS-QoSOnto: A QoS ontology for Web services," in *Proceedings of the 4th IEEE International Symposium on Service-Oriented System Engineering, SOSE 2008*. IEEE, dec 2008, pp. 233–238. [Online]. Available: <http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=4730492>
- [27] Z. Zheng and M. R. Lyu, *QoS Management of Web Services* (Google eBook). Springer, 2013. [Online]. Available: <http://books.google.com/books?id=mTJEAAAQBAJ&pgis=1>
- [28] D. Bianculli, W. Binder, M. L. Drago, and C. Ghezzi, "ReMan: A pro-active reputation management infrastructure for composite Web services," in *2009 IEEE 31st International Conference on Software Engineering*. IEEE, may 2009, pp. 623–626. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1555001.1555092>
- [29] P. Wang, K.-M. Chao, C.-C. Lo, R. Farmer, and P.-T. Kuo, "A Reputation-Based Service Selection Scheme," in *2009 IEEE International Conference on e-Business Engineering*. IEEE, 2009, pp. 501–506. [Online]. Available: <http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=5342065>

M-Learning for Blind Students Using Touch Screen Mobile Apps Case Study- Special Education in Hail

Abdelrahman H. Hussein¹, Majed M. AlHaisoni¹, Ashraf A. Bany Mohammed², Mohammed Fakrudeen³
¹College of Computer Sciences and Engineering, University of Hail, Hail, Kingdom of Saudi Arabia

²School of Business, The University of Jordan, Amman, Jordan

³Dept. of Computing and Technology, Anglia Ruskin University, Chelmsford, United Kingdom

Abstract- The relative newness of the touch-screen (TS) based device creates a phenomenon unique and unstudied in the academic environment with regard to blind students dependent on Braille. This qualitative research study explores how the use of a multi-modal touch-screen based device affects the academic environment for totally blind students using YouTube videos. The pilot program (android app) included a prototype for the English course offered to fifth grade level pupils attending primary school in Hail, KSA. Data collected from students through a survey and focus group interviews and from the faculty through individual interviews was coded and organized according to the research questions. Findings analysis was organized by way of the study's conceptual framework: (a) substitution of Braille course materials with YouTube video lessons (b) accessibility and usability of the developed prototype. Findings concluded that the majority of students in this study perceived YouTube course materials on an touch-screen based device (using android app) to be as good as, or better, than Braille course materials, the multi-modal functionality of the touch-screen based device augmented personal study and classroom learning, and the personal use positively contributed to academic use of the device.

Keywords- Accessibility; Usability; Touch screen; M-learning; YouTube videos; and Blind students

I. INTRODUCTION

Over the centuries, major shifts in literary technology have circled around usability, durability and ease of reproduction. Blind users of touch-screen technology necessarily depend on other cues to users who are not visually impaired. The successful merging of these factors could enhance the effective reading for blind students and create the next advance in literary technology. According to World Health Organization (WHO) figures in 'Global Data on Visual Impairments 2010' published by WHO, Prevention of Blindness and Deafness (PBD) [12], "285 million people worldwide (4.24% of the general population) are visually impaired. 39 million are blind and 246 million suffer low vision (severe or moderate visual impairment)." KSA is not listed as a country contributing data though its nearest listed geographical region probably includes it in the Eastern

Mediterranean Region (ERM). "The 21 countries in the Eastern Mediterranean Region were sorted into two clusters of PPP (Purchasing Power Parity). The first included 13 countries classified as LI (Low Income) and LMI (Lower Middle Income), the second 8 countries classified as UMI (Upper Middle Income) and HI. Data from three countries in the LI/LMI cluster and from one in the UMI/HI cluster were available for estimates." For this region with a population of 580.2 million, 12.5% are recorded as blind, 7.6% recorded as having low vision and 8.2% recorded as having visual impairment. As can be surmised from the quote above, these figures include countries with low income and where incidence of blindness is more prevalent. These figures suggest that KSA may well have a lesser incidence than 12.5% general population blindness. Furthermore the percentage population of blind students would most probably fall below the age of 50, a group which represents 20% of the general population but which currently displays incidence of blindness at 82%. So, although the sample may be small, data collection for this age group and their socio-economic status is expected to produce meager numbers. The results of 25 students in this age band in a city with a population of 412,758 (2010 census) might not stand scrutiny for significant results analysis in the wider population, but they provide positive encouragement for further investigation in a limited field, specifically academia.

This study sought to gain a better understanding of the accessibility of the YouTube phenomenon in the academic environment, specifically for blind students. Knowledge gained from this research can provide increased understanding and inform future developments to incorporate a visually-impaired appropriate, readily accessible YouTube lecture video device in school education. Participants in this study were blind pupils from a public high school in Hail, KSA.

The paper has been organized according to the following sections: Section-1 is the introduction to the research. Section 2 and 3 consists of problem statements and research questions Section-4 consists of the background information and related technological terms. Section-5 contains the research methodology describing the simulation scenarios and the

required configuration. Section-6 analyses the results which is then followed by the conclusion

II. PROBLEM STATEMENT

YouTube EDU had grown to include more than 300 colleges and universities and over 65,000 videos of lectures, news, and campus life were freely available for public viewing [4]. These videos comprise only a portion of the content on YouTube with potential educational value. While it is important to find academic uses of YouTube technology, very little is understood about the accessibility of YouTube videos using touch-screen devices for blind students. Most modern pilot programs with YouTube videos have provided anecdotal feedback instead of rigorous academic research about what influence this will have on education. The adoption of the touch-screen device for blind students has not been researched at all in an academic environment.

III. STATEMENT OF PURPOSE AND RESEARCH QUESTIONS

The purpose of this paper was to explore what could be learned from participant experience of the pilot program utilizing YouTube lectures presented on an android touch-screen based device. Better understanding of m-learning functionality in an academic environment may enhance future effort in substituting Braille print materials for digital materials. Thus, the following research questions were addressed:

1. How do video lectures on an touch-screen mobile affect blind students?
2. How does the accessibility of YouTube lectures using mobile touch-screen affect blind students?

To investigate this phenomenon, the researchers conducted a qualitative case study with the students of the pilot program. This research informs faculty, institutions, and mobile app developers on how to improve the device for future adaptation in the academic environment.

IV. LITERATURE REVIEW

The overall convergence of e-book availability and e-reader technology has laid the foundation for the societal shift from print text to digital text. Now that e-readers are cost-reasonable and more content is available, schools are investigating the use of these devices in the academic environment. However, there is a lingering question of whether students will accept the substitution of printed course materials with digital course materials.

Video is now a common form of media on the Web. The growth of online video is beneficial for those who teach and learn online, as access to video on a broad spectrum of topics becomes increasingly available. The exact amount of video currently hosted online is not known, but the Blinkx (<http://www.blinkx.com>) video search website reports having indexed 35 million hours of video.

Video is not only widely available, but also popular among Internet users. A report from Pew Internet & American Life states that 69% of U.S. internet users watch or download video online and 14% have posted videos [6]. Studies of worldwide Internet traffic from Cisco (2010) [2] also suggest rising interest in video, which now accounts for 26.15% of global broadband traffic. Cisco further reports that over one third of the 50 most heavily visited websites are video sites. Internet traffic rankings from Alexa (2010) [1] and comScore (2010) [3] reveal that YouTube is the most highly visited video destination of them all. Since its creation in 2005, YouTube (<http://www.youtube.com>) has gained meteoric popularity as an online video-sharing website. At the end of its first five years of service, YouTube was receiving more than 2 billion views per day (YouTube, 2010) [11] and users were uploading more than 35 hours of video per [10].

Although YouTube was created as a video-sharing service for the everyday user, the potential for educational use has not gone unnoticed. Over time, scores of colleges and universities have established a presence on YouTube by creating their own video-sharing WebPages called YouTube channels. In March 2009, YouTube announced the launch of YouTube EDU (<http://www.youtube.com/edu>), which is an organized collection of YouTube channels produced by college and university partners. At the end of its first year, YouTube EDU had grown to include more than 300 colleges and universities and over 65,000 videos of lectures, news, and campus life were freely available for public viewing [4]. These videos comprise only a portion of the content on YouTube with potential educational value.

The growth of educational video on YouTube runs concurrent with broader trends in educational video viewership, which rose from 22% to 38% between 2007 and 2009 [6]. Those who seek educational video have numerous places to look for content, including websites created specifically for the purpose of disseminating educational video. It is beyond the scope of this article to delve into a detailed analysis of all of the websites offering academic video content, which has been written about elsewhere [5]. However, a few examples serve to illustrate the breadth of educational video websites. The adult academic audience may enjoy video websites such as Academic Earth (<http://academicearth.org>), BigThink (<http://bigthink.com>), Fora.tv (<http://fora.tv>), and TED (<http://www.ted.com>). K-12 teachers or children may prefer educational video sites like Teacher Tube (<http://www.teachertube.com>) or Watch Know (<http://www.watchknow.org>). It should be noted that some crossover occurs between educational video websites and YouTube. For example, Big Think, Fora.tv, TED Talks, and Teacher Tube each maintain YouTube channels in addition to independent educational video websites. Watch Know also pulls content from YouTube and other video websites for embedding and ranking on its main website at <http://www.watchknow.org>.

The potential of YouTube for online education has been examined previously [9]. One of the obvious benefits of using YouTube in online education is that it provides online access

to vast quantities of free public video on a broad spectrum of topics. It is a simple matter to link to or embed YouTube videos in online course content or discussion forums. Content management is also a benefit. Online educators can establish YouTube channels to collect, organize, host, and distribute video. YouTube videos may be grouped into one or more "Video Playlist Lessons," which are created by collecting videos into a playlist then typing a lesson plan into the playlist description area. Playlist lessons have been created to meet learning objectives across the cognitive, affective, or psychomotor learning domains in real-world online classrooms [7]. Online educators may also create interactive video games, simulations, or tutorials by linking videos together through the Annotations tool on YouTube [8].

In July 2010, Snelson (2010a)[7] published the results of a study that compared the speed of reading on different devices. This research measured reading speed on a personal computer, printed book, Amazon Kindle, and the Apple iPad. The participants met all comprehension objectives, but read 6.2% slower than print on the iPad and 10.7% slower on the Kindle. Nielsen ultimately determined the results not statistically significant to conclude that reading is actually slower on an e-reader. However, a surprising finding indicated the participants found the reading of print more relaxing than reading electronic text.

In June 2009, the National Federation of the Blind (NFB) and the American Council of the Blind (ACB) filed a lawsuit against Arizona State University regarding their pilot of the Kindle DX in a college course. The device's inability to be fully used by blind students violated both the federal Americans with Disabilities Act and the Rehabilitation Act of 1973 (Case 2:09). The lawsuit was settled out of court in January 2010, and served as a message that all e-readers, if used in an educational setting, must be accessible by all students. On June 29, 2010, the United States Department of Justice and Department of Education jointly published a letter to college and university Presidents reiterating the results of the lawsuit and mandating compliance in the future.

The researchers also examined different navigation modes for presenting text material: page-by-page (paging) or scrolling. Ninety percent of polled users preferred paging in the portrait layout. E-reading for pleasure was exclusively linear. The table of contents was the most important feature of e-text—followed by hyperlinks, illustrations, page numbers, headings, and highlighted words. On the e-reader itself, users highlighted legibility, portability, easy navigation, ample storage and ease of use as important attributes. 78.9% percent preferred reading on a dedicated e-reader compared to a computer screen. "96% disagree with the statement that the e-reader makes them lose the context of what they read, and more than 70% feel they can both deep read and skim with their e-reader".

V. METHODOLOGY

This section outlines the research methodology that was used in this case study with the following sections: 5.1

rationales for qualitative case study method, 5.2 research questions and propositions, 5.3 data collection methods.

5.1 Rationale for Qualitative Case Study Method

This research explored what could be learned about accessibility to YouTube video lectures from participant experience of the pilot program utilizing an android touch-screen based device. This study was explicitly interested in participant reception derived from the experience using video lectures on a touch-screen device, with a specific application within the social context of the pilot program. Therefore, qualitative research was an appropriate match for the intent of this study.

The context of the case included the pilot program faculty and pilot program student experience of using video lectures presented within touch-screen mobile using android. Holistically, the subject of this study was not the participants themselves, but the participants' perceptions of using video lectures using a touch-screen mobile device and the accessibility in dealing with it.

5.2 Research Questions and Propositions

The study propositions, also known as the theoretical framework of a qualitative study, were derived from the literature review (Merriam, 1998)[13]. The research questions with their ancillary propositional questions were as follows:

Q1. How do students perceive video lectures on touch- screen based mobile?

1.1 How does the replacement of Braille course materials with video lectures, regarding frequency of listening, affect students?

1.2 How does the replacement of Braille course materials with video lectures, regarding duration of reading, affect students?

1.3 How does the replacement of Braille course materials with video lectures, regarding speed of reading, affect students?

1.4 How does the replacement of Braille course materials with video lectures, regarding understanding of course materials, affect students?

1.5 How does the replacement of Braille course materials with video lectures, regarding class participation, affect students?

Q2. How does the accessibility of YouTube lectures using touch-screen mobile affect students?

To answer these questions, the following feedback was taken using Likert type scale (1-Strongly, Disagree, 5-Strongly Agree) questionnaire with respect to the original prototype developed by The University of Hail / College of Computer Science and Engineering (Fig. 1).

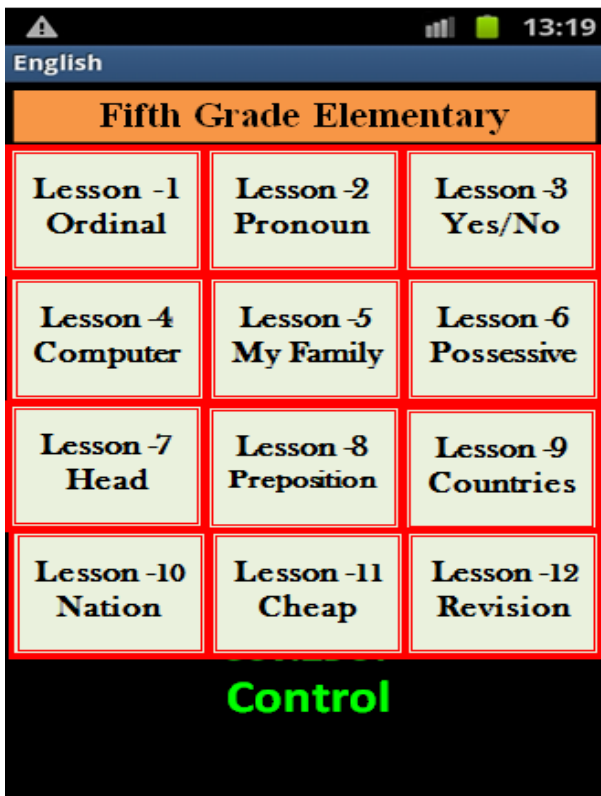


Fig. 1 Prototype Design

5.3 Data Collection Methods

All methods of data collection may be used in case study research (Creswell, 2009; Merriam, 1998; Yin, 2009). In this paper, multiple methods of data collection were utilized in researching the Samsung Galaxy Touch-screen at the Public School, Hail. Data collection for this study occurred between 4th January 2015 and 4th February 2015. The survey and focus group methods were used for blind student participants. Research was divided into three sections. Firstly, a pilot test was conducted using an original prototype developed by University of Hail. Secondly, blind students completed questionnaires to satisfy the first research question. Thirdly, feedback based on the prototype's accessibility, was taken.

The survey method was the best fit to collect data about the blind student feedback at school level. The survey collection tool was selected based on the survey's strength of being unobtrusive, while allowing the researcher to collect information directly from a large group of participants. Typically considered a quantitative tool, in this study a survey was used to build a more complete picture of the processes and perceptions of students' experiences with the digital course materials. The survey used both closed and open-ended questions.

These open-ended questions helped provide insight into the participants' experience and perceptions. The quantitative data collected by the survey is presented and analyzed utilizing descriptive statistics to inform the qualitative study.

The student survey instrument was developed directly from the primary research questions and their supporting propositions. Each survey question was tied directly to one or more proposition.

VI. RESULTS AND DISCUSSION

The following is a presentation of the findings. The first research question is supported by multiple propositional questions. These propositions act as a guide to inform the research questions from across the research tools. Each proposition will be given a preliminary finding, which will then be collected into an overall finding for the first research question.

Research question 1. *How do students perceive video lectures on touch-screen based mobile?*

The first research question is supported by five research propositions. Each of the propositions will be presented with supporting findings from the research followed by an overall finding for the research question.

Proposition 1.1 *How does the replacement of Braille course materials with YouTube video lectures, regarding frequency of listening, affect students?*

The research asked respondents if they listen more or less often when using the prototype than Braille course materials and why they believed they were listening more or less often. 12 out of 25 blind participants felt they were listening more often than with the Braille materials as the novelty of discovering new ways of reading appealed to them and they were easier to learn than touching the Braille letters. 5 out of 25 respondents felt they listened less often, with two giving the reason that they worked in a secure facility in which they were not allowed to take the TS. The other respondents (8 of 25) felt reluctance at exploring the device, lacking confidence with something new, and that they needed time to familiarize themselves to TS before they could adapt to this culture.

Finding 1.1 *48% of students perceived their frequency of listening to YouTube videos to be about the same or more often due to portability (Fig. 2).*

Proposition 1.2 *How does the replacement of Braille course materials with YouTube video lectures, regarding duration of reading, affect students?*

Respondents were asked if they listened for longer or shorter periods of time when using the TS. A majority of respondents (17 of 25) felt their duration of listening to YouTube videos in TS was much shorter than actual lectures in the classroom. Some (7 of 25) felt it took a longer time to listen to YouTube lecture videos than the class room where they can clear doubts instantly rather than having to replay YouTube video lectures for clarification or understanding. Only a minority of student focus groups responded that they did not experience a change in their preparation for class.

Finding 1.2 The majority (68%) of students felt their duration of listening to course materials was shorter than when reading using the Braille course material (Fig. 3).

Proposition 1.3 How does the replacement of Braille course materials with YouTube video lectures, regarding reading speed, affect students?

Participants were asked if they read or listened more quickly or less quickly when using the TS (YouTube lectures). 64% (16 of 25) of the respondents listened easily with TS while 12% listened with difficulty and 24% didn't feel any difference.

Finding 1.3 64% felt more comfortable listening to YouTube lectures over reading Braille material (Fig. 4).

Proposition 1.4 How does the replacement of Braille course materials with YouTube video lectures, regarding understanding of course materials, affect students?

The research asked participants if they found they understood more or less of what they read when using the Braille material. Of the respondents the clear majority 56% (14 of 25) understood more, 28% (7 out of 25) felt that they understood less and the remaining 16% (4 out of 25) felt no difference between the two.

Finding 1.4 The clear majority (56%) of students stated understanding is more readily appreciated in YouTube lectures as they can repeat the YouTube lectures more often than Braille course materials, which is time consuming (Fig. 5).

Proposition 1.5 How does the replacement of Braille course materials with YouTube video lectures, regarding class participation, affect students?

The research asked respondents if they found themselves participating more or less in class after having read YouTube lecture videos in TS. A majority 52% (13 of 25) of the respondents found themselves participating about the same. 28% (7 out of 25) participated more and another 20% (5 out of 25) participated less.

Finding 1.5 The overwhelming majority (52%) of students participated the same amount or more after having YouTube lecture videos on TS (Fig. 6).

Research question 2. How do blind students rate the accessibility of YouTube lectures using touch-screen mobile?

The second research question is supported by feedback of usability features with Likert Scale. Each of the usability features will be presented with supporting findings from the research followed by an overall finding for the research question.

The research reveals that almost all blind users were satisfied with the accessibility features applied to the YouTube lecture videos. Fig. 7 shows the average of the Likert scale used against the usability features. Most of the usability features were above 4. The last usability feature asked the group if they required any more accessibility features for the prototype. Most of the blind participants felt

that all accessibility features were met and needed no further improvement. The overall rating shows they are very much satisfied with the prototype.

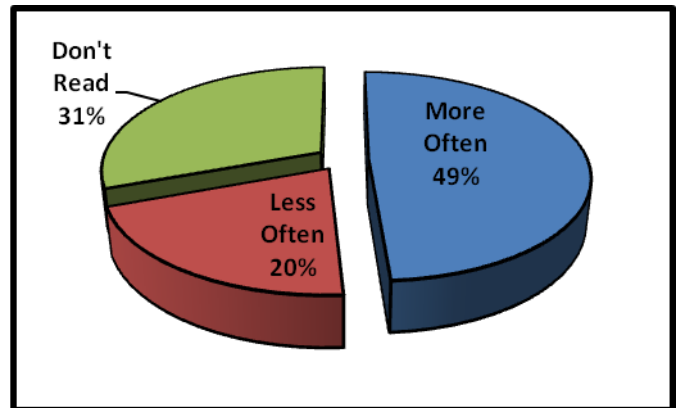


Fig. 2 Frequency of Listing in YouTube lecture

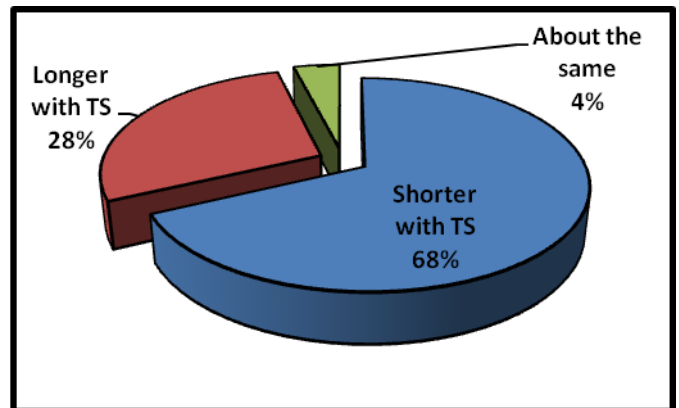


Fig. 3 Duration of reading in YouTube lecture

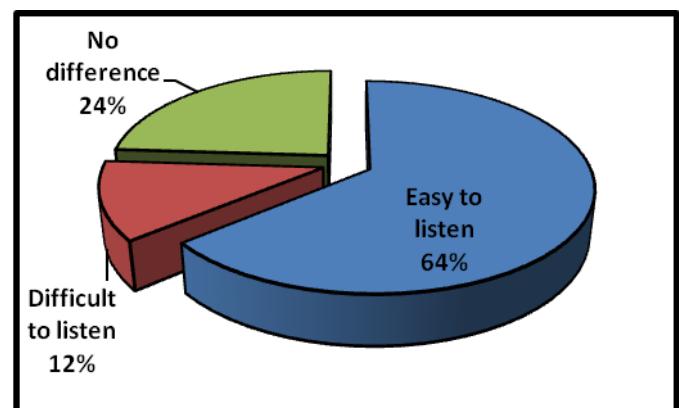


Fig. 4 Speed of reading in YouTube lecture video

VII. CONCLUSION

The purpose of this study was to explore what could be learned from participant experience in a pilot program in regard to replacing Braille course materials with YouTube lectures presented via touch-screen based mobile device. The conclusions of this study follow the research questions, findings, and analysis, and address the following four areas: (a) usability of the touch-screen based mobile device and, (b) the touch-screen based mobile device enhanced learning.

The touch-screen based mobile device enhanced learning.

The first major finding of this research is that the majority of students found that reading course materials on an touch-screen based mobile device did affect their reading duration, speed of reading, reading comprehension, and class participation. It can be assumed that time plays a large part in the learning experience. If something takes a long time to do, such as study reading, students can become fatigued, bored and distracted, or find they have other demands on their time and their study session must end prematurely, perhaps before they have had time to grasp a new concept. By providing technology to overcome these negative aspects of 'lost time', more can be learned in terms of quantity and perhaps quality too when students are motivated to find out more. Repetition of good technique is essential to skill acquisition, as much to operating a touch-screen until it is second nature to the user as it is to listen to an audio book and hear well-constructed sentences, appreciate a scientific argument or construct empathy with novel characters and expand vocabulary beyond a student's current reading age. Mainstream schools are utilizing technology in the classroom to give their students rich experience and to make learning fun. Blind students using touch-screen technology can become part of that mainstream environment more readily if they share similar experiences with colleagues. Two of the elements suggested to improve cognitive function are to develop social connections and curiosity and creativity. This prototype may work towards providing those elements for blind students and to make their classrooms more inclusive.

Usability of the touch-screen based mobile device.

The second and third major findings of this study expressed how the multi-functionality of the touch-screen based mobile device is accessible for blind students. The majority of students perceived the touch-screen based mobile device as a easy to use, easy to navigate, felt in control, intuitive and easy to learn. However, for academic purposes, the touch-screen based mobile device was primarily used as a content consumption device in conjunction with a personal computer. The third finding stated that a clear majority of students found the touch-screen based mobile device personally useful, carried it with them more often, and found themselves using it more academically due to its convenience and portability. Therefore, this study also concludes for blind students in this study, the portability and accessibility of the touch-screen based mobile device contributed positively to academic use of the device.

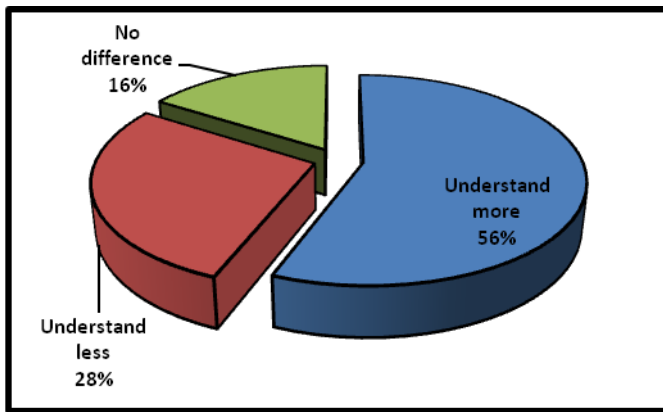


Fig. 6 Understanding of lectures in YouTube versus class room

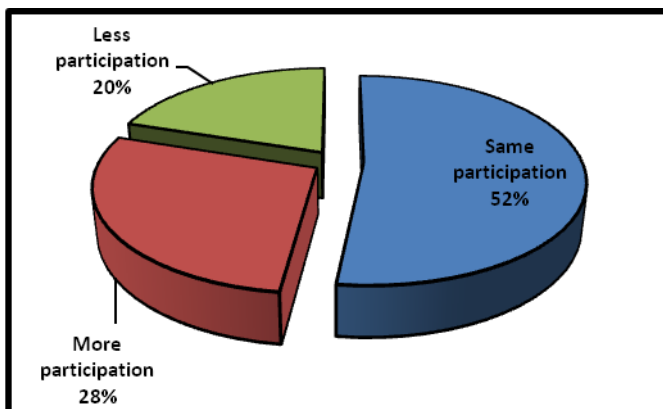


Fig. 7 Class participation after listening to YouTube lecture

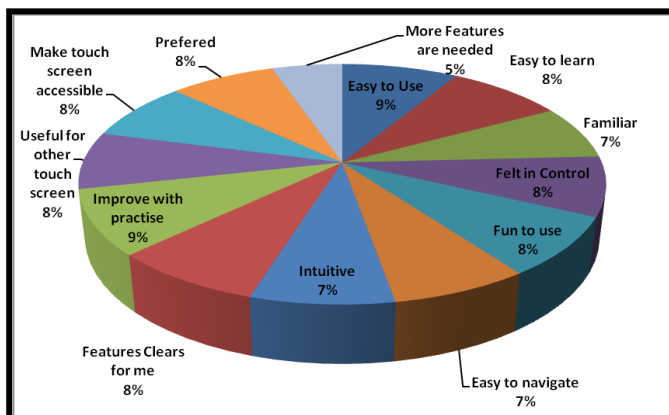


Fig. 5 Usability feature analysis for YouTube prototype

ACKNOWLEDGMENT

The authors thank Research Deanship in University of Hail, Saudi Arabia for its financial support

REFERENCES

- [1] Alexa (2010). *YouTube.com site info*. Retrieved from <http://www.alexacom/siteinfo/youtube.com>
- [2] Cisco (2010). *Cisco visual networking index: Usage study*. Retrieved from http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/Cisco_VNI_Usage_WP.html
- [3] ComScore (2010, May). *comScore Releases May 2010 U.S. online video rankings: YouTube streams all-time high of 14.6 billion videos viewed*. Retrieved from http://www.comscore.com/Press_Events/Press_Releases/2010/6/comScore_Releases_May_2010_U.S._Online_Video_Rankings
- [4] Greenberg, O. (2010, March 25). More courses, more colleges: YouTube EDU turns one. *The Official YouTube Blog*. Retrieved from <http://youtube-global.blogspot.com/2010/03/more-courses-and-more-colleges-youtube.html>
- [5] Miller, M. (2009). Integrating online multimedia into college course and classroom: With application to the social sciences. *MERLOT Journal of Online Learning and Teaching*, 5, (2). Retrieved from http://jolt.merlot.org/vol5no2/miller_0609.htm
- [6] Purcell, K. (2010). *The state of online video*. Retrieved from the Pew Internet & American Life Project website: <http://pewinternet.org/Reports/2010/State-of-Online-Video.aspx>
- [7] Snelson, C. (2010a). Mapping YouTube "video playlist lessons" to the learning domains: Planning for cognitive, affective, and psychomotor learning. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2010* (pp. 1193-1198). Chesapeake, VA: Association for the Advancement of Computing in Education. Retrieved from <http://www.editlib.org/p/33518>
- [8] Snelson, C. (2010b). Virtual movie sets and branching video: Developing interactive educational machinima with Second Life and YouTube. *Proceedings of the Technology, Colleges, and Community Worldwide Online Conference*. Retrieved from <http://etec.hawaii.edu/proceedings/2010/Snelson.pdf>
- [9] Snelson, C. (2009). Web-based video for e-Learning: Tapping into the YouTube phenomenon. In H. Yang and S. Yuen (Eds.), *Collective Intelligence and E-Learning 2.0: Implications of Web-Based Communities and Networking* (pp.147-166). Hershey, PA: IGI Global.
- [10] Walk, Hunter, (2010, November 10). Great Scott! Over 35 hours of video uploaded to YouTube every minute. *The Official YouTube Blog*. Retrieved from <http://youtube-global.blogspot.com/2010/11/great-scott-over-35-hours-of-video.html>
- [11] YouTube (2010, May 16). At five years, two billion views per day and counting. *The Official YouTube Blog*. Retrieved from <http://youtube-global.blogspot.com/2010/05/at-five-years-two-billion-views-per-day.html>
- [12] <http://www.who.int/blindness/GLOBALDATAFINALforweb.pdf?ua=1>
- [13] Merriam, S. B., & Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass Publishers

A Fitness-Gossip Routing Protocol for Saving Energy in Wireless Sensor Networks

Kareem Radi Hassan

Department of Computer Science, University of Basrah, Basrah, Iraq

Abstract: Gossiping is traditional routing scheme commonly used in WSNs to transport data in a multi hop manner because of limited radio range and energy constraints. Gossiping protocol is simple and do not involve additional devices, but at the same time it faces several deficiencies when it used in WSNs. This paper describe an efficient technique to transmit the message to the sink node. The main idea behind the new protocol which is called (Fitness-Gossiping Protocol FGP) which is a modification of Gossiping protocol by using a fitness function to select the optimum next node, when the optimum next node is selected the data is transmitted to the next node. We discussed how the new approach saved the energy of the network and to achieve maximize the network lifetime in comparison with its counterparts. In the same time, the Fitness-Gossiping protocol provides the balanced energy between nodes.

Index Terms— Gossiping, Fitness- Gossip, network lifetime, routing, wireless sensor networks (WSNs).

1. INTRODUCTION

New advances in wireless communications leads to introduce the wireless sensor networks (WSNs) and to provide low-cost networking, low power consumption, and tiny devices, often for low data rate transmission. Self-organizing capabilities are required in WSNs due to the random deployment in ubiquitous environments that cannot be accessible in this environments [1]. WSNs are designed for specific application, the

application requirements are changed according to the nature of the application. Hence, the requirements of routing protocols are changed depending on the nature of the application. For instance, the requirements of routing protocols designed for environmental applications are different in many aspects from those designed for military or health applications [2-4]. However, regardless of the application, routing protocols must attempt to maximize the network life time and minimize the overall energy consumption in the network. Network lifetime is a critical aspect in the build of WSNs. In many applications, replacing or recharging sensors is some- times impossible [5].

Unbalanced power exhaustion is an impact problem in WSNs characterized by the multi-hop routing and many-to-one traffic pattern. This uneven power waste can significantly minimize network lifetime. Usually in routing algorithm, the best path is chosen for sending of data from origin to destination. Over an interval of time, if the same route is chosen for all transmissions in order to obtain battery performance in terms of quick transmission time, then those nodes on this route will consume energy faster than the other nodes in network [6-8].

In this paper, the proposed method seeks to investigate the problems of balancing energy consumption and maximization of network lifetime for WSNs. We propose a new approach by adding fitness function that depending on power ratio, the distance of next node from the source node, and the distance of the next node from the sink. Then select the

optimal routing path from the source to the destination.

The rest of this paper is organized as follows. Background and related work and related concepts of the routing algorithm to maximize the WSN lifetime is presented in Part 2. In Part 3, the paper describes the new technique. Simulation and evaluation is presented in Part 4. Finally, conclusion and discussion are presented in Part 5.

2. BACKGROUND

Gossiping protocol uses a elementary forwarding rule and does not need costly topology servicing or complex path discovery algorithms [9]. The idea of gossiping protocol is that each sensor node should send the received packet to a randomly selected neighbor. Upon receiving the packet, the neighbor chosen randomly selects one of its own neighbors and transmits the packet to the neighbor chosen; this can include the node which sent the packet. This operation is repeated in each node until the packet reaches its intended destination. Gossiping solves the implosion issue by limiting the number of packets that each node forwards to its neighbor to one copy. Gossiping protocol may suffers from the excessive latency due to the random nature of sending packet from source to its destination. Gossiping is not a reliable method for data dissemination, because some nodes in the large network may not receive the message at all since the neighbor is selected randomly. The main disadvantage of this protocol is that the packet delay is very long [10]. Figure (1) and Figure (2) illustrate two examples of Gossiping protocol [11].

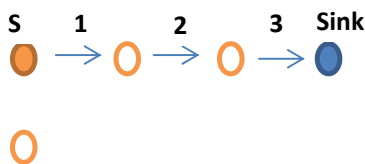


Fig. 1. Gossiping (scenario 1).

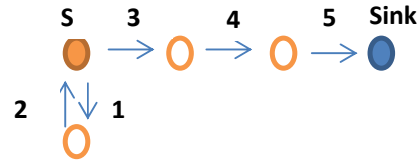


Fig. 2. Gossiping (scenario 2).

Flooding protocol consumes higher energy than Gossiping. However, Gossiping have many problems such as latency; data transmission slowly. Despite the simplicity and inefficiency of Flooding and Gossiping, they could be used for specific functions, for example, during deployment phases and network initialization [12, 13]. The power consumed by Gossiping, is approximately equal to $O(K^L)$.

K : number of nodes that forward the packet.

L : number of hops before the forwarding stops.

The most noteworthy feature of Gossiping is the capability to control the energy exhaustion by appropriately selecting K and L [14].

After a technical review of Gossiping protocol we can determine these disadvantages:

- The next hop is randomly selected, it may include the source itself.
- The packet will travel through these chosen neighbors until it reaches the sink or override the number of hops.
- It suffers from packet loss.

The most significant issue of Gossiping is that it suffers from latency caused by data propagation.

The lifetime issue in WSNs has received significant attention by the researchers. there are many related work has been proposed during the last years such as; the work in [15]

proposed to minimize the hop stretch of a routing path (defined as the ratio of the hop distance of a given path to that of the shortest path) in order to reduce the energy cost of end-to-end transmission. The approaches in [16], [17] took a different view for prolonging the network-lifetime. They attempt to sustain the availability of the sensors that have less energy by distributing the traffic load to the ones with much residual energy. All of the above-mentioned works focus on improving energy-efficiency using fixed routing paths; nonetheless, due to the lack of path diversity, those nodes traversed by fixed routing paths may drain out their energy quickly. The work in [18] exploited two natural advantages of opportunistic routing, i.e. path diversity and the improvement of transmission reliability, to develop a distributed routing scheme for prolonging the network lifetime of a WSN.

The goal of this work is to assist each sensor in determining a suitable set of forwarders as well as their priorities, thus, enabling effort to extend the network-lifetime.

3. The proposed technique

With some changes to the Gossiping protocol we can decrease the energy consumption and also increase the network lifetime. Therefore, in order to resolve the drawbacks of the Gossiping protocol, we have proposed a new protocol as an extension for Gossiping. In this protocol we have increased the network lifetime by selecting a node with a maximum residual energy and lower distance to the sink. We have also achieved a high packet delivery ratio and reduced the delay in delivering the packet. A number of different metrics have been used to prolong the lifetime of the sensor networks. These metrics are as follows:

- 1) *Remaining Energy (RE)*: The algorithm calculate the remaining energy of the next node as a ratio to the initial energy for all sensors.

- 2) *Next Node Distance (NND)*: is a ratio of the distance of the next node to the range of the transmission signal of the sensor.
- 3) *Sink Distance (SD)*: is the ratio of the distance of the sink from the next node.

These metrics can be calculated as a fitness function for the selected of the next node from the current node as in Equation (1) and Figure (3):

$$\text{Fitness} = (\text{RE}/\text{E}) + (\text{NND}/\text{Range}) + (1 - \text{SD}/\text{Sink Distance}) \dots\dots\dots (1)$$

Where RE is the Remaining Energy in node, E is the initial Energy of the node, NND is the Next Node Distance from the current node, Range is Limit of transmission distance, SD is the ratio of the distance of the sink from the next node.

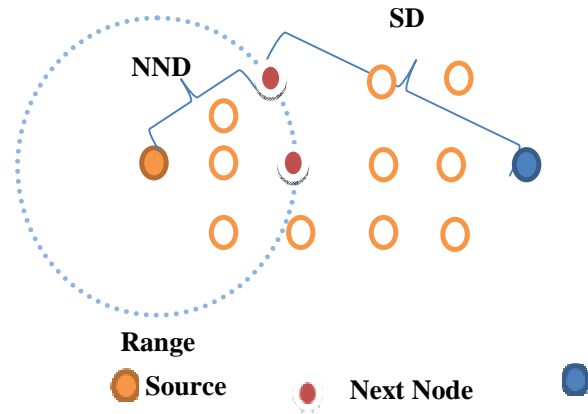


Fig. 3. The method metrics

1. The fitness function calculates the fitness for all nodes neighbors to the current node within the range of the source node, calculates the remaining power for all neighbors, the distances of neighbors, and the distances of neighbor for the sink. Then select the highest fitness to pass the packet for it. Repeats these steps until reaching the sink.

4. Simulation and Evaluation

In simulation 100 sensor nodes are randomly deployed in a topographical area of dimension 100 m × 100 m. topographical area have the sensed transmission limit of 25 m. The performance of the proposed method is tested in this topographical area. There is only one data sink which located at (90 m, 90 m). All sensor nodes have the same initial energy 0.5J. Transmission and receiving costs are characterized by the expressions:

$$EnT(k) = Eelec \cdot k + Eamp \cdot k \cdot dl \dots\dots(2)$$

And

$$EnR(k) = Eelec \cdot k \dots\dots\dots(3)$$

where k is the number of bit per packet, dl is the distance from the sender node to the receiver node, $Eelec$ and $Eamp$ are per bit energy dissipation in transmitting or receiving circuitry and energy required per bit per meter square for the amplifier to achieve acceptable signal to noise ratio (SNR) respectively. Simulations are done using the values 50 nJ/bit and 100 pJ/bit/m² for $Eelec$ and $Eamp$, respectively.

TABLE I
SIMULATION PARAMETERS

| Parameter | Value |
|---|---------------------------|
| Topographical Area (meters) | 100 m × 100 m |
| Sink location (meters) | (90, 90) |
| Number of nodes | 100 |
| Limit of transmission distance (meters) | 25 |
| Initial energy of node | 0.5 J |
| Eelec | 50 nJ/bit |
| Eamp | 100 pJ/bit/m ² |
| Packet data size | 100 bit |
| Number of transmission packets | 2 × 10 ⁴ |

We assume some parameters to implement the simulation results:

Range: the range of coverage of the sensors is 25 m. Any sensor that detects the event

draws the transmission range to limit the number of nodes in its transmission range.

Remaining Energy: shows the amount of energy remaining. We assume that in the beginning the energy of all nodes is same.

Location of sensor: all sensors have GPS or other location devices and can move to any position (with known coordinates) within their mobility range. It can be shown that the proposed Fitness-Gossiping routing protocol performs better when compared to the other routing protocols. We investigated the performance of the proposed protocol by comparing packet loss, delay, live node and total energy saving per round.

Number of hops: as seen in Figure 4, number of hops is at a minimum in Fitness-Gossiping when compared with Gossiping. However, the number of hops increased after the 1500th iteration due to the Fitness-Gossip search on the better nodes with high energy and very close to the sink.

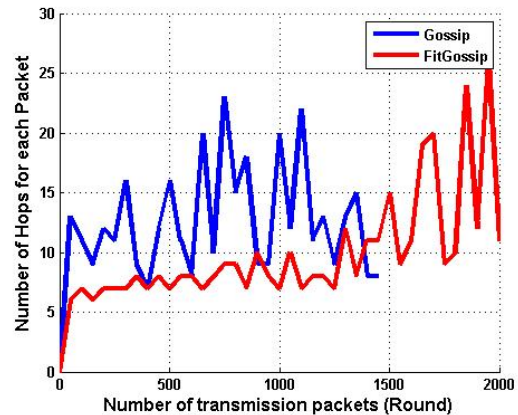


Fig. 4. Number of hops for gossiping and Fitness-Gossip

Live nodes: after the 250th iteration in Gossiping the nodes start die. However, as shown in Figure 5, in our proposed protocol the nodes start die after approximately 1500 iterations by balancing and using the energy in a fair manner.

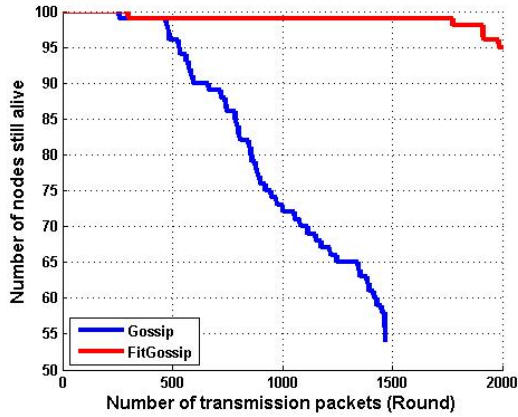


Fig. 5. The number of nodes still alive in Gossiping and Fitness-Gossiping.

Delay: in contrast to other protocol that randomly select the next hop for the packet, in our protocol the fittest node that have the best fitness is selected as the next hop. Therefore the delay in our protocol is the smallest among the protocol compared. We can see this clearly from the simulation results in Figure 6. Until the 450th iteration the delay is fixed approximately at 2msec.

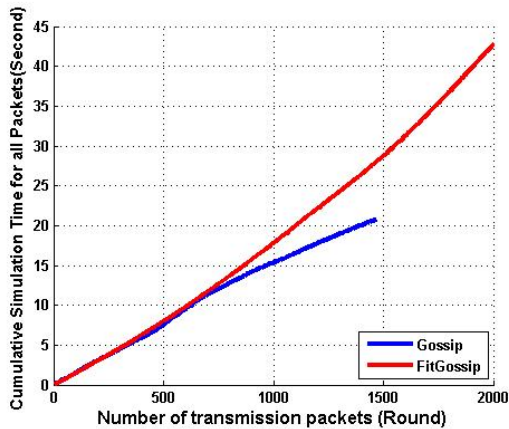


Fig. 6. Cumulative simulation time for Gossiping and Fitness-Gossiping.

Energy Consumption: in our proposed protocol the relay nodes are not selected blindly (without knowing their fitness) as is done in the other routing protocol in this comparison. Moreover, energy reduction for each node occurs for every transmission or

reception made. Hence, the probability of choosing the same node as the next hop is reduced. Thereby, the energy has been balanced and fairly used. All this leads to saving energy and hence prolonging the overall network lifetime compared to the other protocols, as seen in Figure 7.

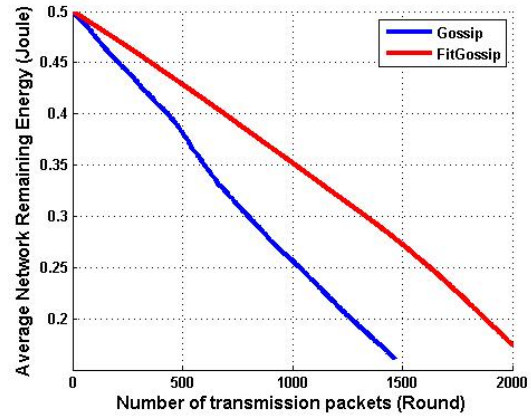


Fig. 7. Average remaining energy for Gossiping and Fitness-Gossiping.

5. Conclusion

The nodes in wireless sensor networks operates on finite battery power, this issue leads make the effective use of power is very important. The path selection can effect on network lifetime. These issues needs to balance in power consumption to increase network lifetime. To enhance the network lifetime by decrease the power consumption by select the best path from start node to sink, we proposed a new method that can provide this goals by using a fitness function to calculate the pest path using the remaining energy and the distance from the sink. The technique provide highest remaining energy, minimum number of hops, highest number of nodes and low delay. The evaluation of the method is compared with the Gossiping protocol to show the difference using the same criteria.

References:

- [1] J. Yick, B. Mukherjee, and D. Ghosal. Wireless Sensor Network Survey. *Elsevier Computer Networks*, 52(12): 2293-2330, 2008.
- [2] Á. Lédeczi, A. Nádas, P. Völgyesi, G. Balogh, B. Kusy, J. Sallai, G. Pap, S. Dóra, K. Molnár, M. Maróti and G. Simon, "Counter sniper System for Urban Warfare," *ACM Transactions on Sensor Networks*, Vol. 1, No. 2, 2005, pp. 153-177. doi:10.1145/1105688.1105689
- [3] T. He, S. Krishnamurthy, J. A. Stankovic, T. Abdelzaher, L. Luo, R. Stoleru, T. Yan, L. Gu, G. Zhou, J. Hui and B. Krogh, "VigilNet: An Integrated Sensor Network System for Energy-Efficient Surveillance," *ACM Transactions on Sensor Networks*, Vol. 2, No. 1, 2006, pp. 1-38. doi:10.1145/1138127.1138128
- [4] R. Verdone, D. Dardari, G. Mazzini and A. Conti, (2007). "Wireless Sensor and Actuator Networks Technology, Analysis and Design," 1st Edition, Elsevier, London, 2007.
- [5] S. Bandyopadhyay and E. Coyle, "An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks," *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003)*, San Francisco, 30 March-3 April 2003, pp. 1713-1723.
- [6] H. Zhang and H. Shen, "Balancing energy consumption to maximize network lifetime in data-gathering sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 20, no. 10, pp. 1526-1539, Oct. 2009.
- [7] K. Akkaya and M. Younis, "A survey of routing protocols in wireless sensor networks," *Ad Hoc Netw.*, vol. 3, no. 3, pp. 325-349, May 2005.
- [8] F. Ren, J. Zhang, T. He, C. Lin, and S. K. Das, "EBRP: Energy-balanced routing protocol for data gathering in wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 12, pp. 2108-2125, Dec. 2011.
- [9] Kemal Akkaya and Mohamed Younis, "A Survey on Routing Protocols For Wireless Sensor Networks" *Ad Hoc Networks*, Volume: 3, Issue: 3, May 2005.
- [10] Yazeed Al-Obaisat and Robin Braun, "On Wireless Sensor Networks: Architectures, Protocols, Applications, and Management". University of Technology, Sydney, Australia, 2007.
- [11] Ghaihab Hassan Addy, "Gossiping with Route Poisoning: New Routing Strategy in Wireless Sensor Networks", M.Sc. thesis, University of Basrah, Iraq, 2011.
- [12] W. Rabiner Heinzelman, J. Kulik and H. Balakrishnan, "Adaptive Protocols for Information Dissemination in Wireless Sensor Networks," *Proceedings of the Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '99)*, Seattle, Washington, 15-20 August 1999, pp. 174-185.
- [13] S. M. Hedetniemi, S. T. Hedetniemi and A. L. Liestman, "A Survey of Gossiping and Broadcasting in Communication Networks," *Networks*, Vol. 18, No. 4, 1988, pp. 319-349. doi:10.1002/net.3230180406
- [14] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," *IEEE Transactions on Wireless Communications*, Vol. 1, No. 4, 2002, pp. 660-670. doi:10.1109/TWC.2002.804190
- [15] M. J. Tsai, H. Y. Yang, and W. Q. Huang, "Axis-based virtual coordinate assignment protocol and delivery-guaranteed routing protocol in wireless sensor networks," in *Proc. IEEE INFOCOM 26th Int. Conf. Comput. Commun.*, May 2007, pp. 2234-2242.
- [16] J. Park and S. Sahni, "An online heuristic for maximum lifetime routing in wireless sensor networks," *IEEE Trans. Comput.*, vol. 55, no. 8, pp. 1048-1056, Aug. 2006.
- [17] C. Wu, R. Yuan, and H. Zhou, "A novel load balanced and lifetime maximization routing protocol in wireless sensor networks," in *Proc. IEEE Vehicular*

Technol. Conf. VTC Spring, May 2008,
pp. 113–117.

- [18] C. C. Hung, K. C. J. Lin, C.-C. Hsu, C.-F. Chou, and C.-J. Tu, “On enhancing network-lifetime using opportunistic routing in wireless sensor networks,” in *Proc. 19th Int. Conf. Comput. Commun. Netw.*, Aug. 2010, pp. 1–6.

Computing the dynamic reconfiguration time in component based real-time software

Asghar farhadi

Mehran sharafi

Computer depart

Computer depart

Islamic azad university Arak branch

Islamic azad university Najaf abad Branch

Arak, Iran

Najafabad ,Iran

Abstract

New microcontrollers with enhanced application capabilities for more complex scenarios are developed. However, the settings are volatile and ever-changing environment requires permanent systems that compatible with the new conditions.

Dynamic re-configuration has a powerful mechanism for the implementation of the adaptation strategy. Real-time control system is one of the challenges for the implementation of dynamic re-configuration software.

In previous works the Adapt.NET is adapted in framework of the implementation of component-based applications. A new web-based test complete re-configuration is proposed here with a limited time. The application dynamically re-configures the client-side component compatibility in case of failure of the component parts.

in this article the timing behavior of the implemented dynamic reconfiguration algorithm is analyzed. The manner of the hybrid component-based applications adaptation during environmental condition changes are described as well.

In this article ,in order to predict the implementation time the behavior of dynamic reconfiguration algorithm and the manner of real-time planning that can be adapted to environmental changes are assessed, as well as the correlation of reconfiguration during the deadline period.

Key words: dynamic reconfiguration, blackout, reconfiguration time, adaptation, state

I INTRODUCTION

Embedded systems constantly face environmental conditions' challenges. A system to protect the quality parameters like changes in energy consumption and change in the rate of consumption of communication bandwidth memory should be developed.

The available embedded systems are reconfigured on an off line basis.

Stop running process are costly and can be detrimental for the users

Andreas Rasche et al introduce the Adapt.NET framework which consists of tool for building distributed applications based on components and a tool for monitoring the runtime application.

If the adjustment program is the program to be subjected to downloading, a new version of the program

the components and connections between/among them is named configure program.

After of implementation the new version of the infrastructure the dynamic structure is developed. Developers can choose adaptable programs from a variety of architectural styles [1].

The Adapt.NET framework applied in this study is of distributed control laboratory type, which has an open infrastructure testing an open infrastructure for testing the controlling robots through in the internet [4].

In lab, the accuracy of laboratory testing with the use of redundancy and dynamic reconfiguration is assured and the result, are analyzed.

The malicious codes downloaded from the internet can damage the hardware or cause a change running codes therefore, the observers check the components, User behavior and if an error is detected by the control components the program will be re-configured to correct the occurred error.

The conducted test in this study is run through a pc x86 and the operating system Windows 7

Dynamically re-configured new version should be reconfigured in limited time.

Dynamic reconfiguration is a synonym of runtime evolution, the ability to change a system's functionality and/or topology while it is running. Dynamic reconfiguration typically involves changes at runtime via addition, deletion, and replacement of components, and/or alteration of the topology of a component system. The benefit is application consistency and service continuity for the system while it is being updated [17].

In this article how the dynamic configuration and how the program reconfigures in real-time is explained.

The rest of the article is as follows: Section 2 shows describes Adapt.NET framework for compatibility software distributed; consequently, a general model for computing applications is given

II METHODOLOGY AND MATERIAL

The method adopted in this study is descriptive analytic, here, an interface framework for dynamic re-configuration is introduced. The Component and connector through XML description language is described [2].

An overview of the infrastructure re-configured dynamically named is CoFRA is shown in fig 1.

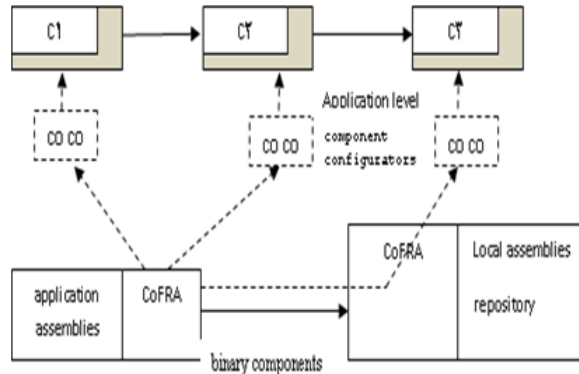


Figure1. Reconfiguration[3]

When a program is run, the CoFRA reviews the transmitting net components through the CoFRA with respect to performance sustainability within the network.

The dependent components are developed by a central component repository.

The implementer version features will be able to update each component that must be reconfigured like methods for connects setting component parameters and the transfer of components.

The tool applied to produce this interface configuration is shown in cream, Fig (1).

Components can be introduced in different manners.

These components can be on a machine or a remote component.

The hidden components complexity, when accessing different parts of component type are illustrated in Fig (1).

For example, if a component is run by different processes it regulates the communication channel interface components, the implementation of which is based on .NET framework.

In this newly introduced network java or CORBA components can be applied. The correlations among .NET, CORBA and Java through IIOP.NET are introduced in [16].

Framework of other unwanted components can be integrated easily.

In general, applications for adjustment of the real need for a strategy in adapting the application time, the properties will be reconfigured according to their environ state.

There exist three types of environ parameters:

- Condition: outside the parameters of the program

- State of the software components: e.g. components of the crash can be determined by a variable and
- Attributes of components: to monitor the of the software, the internal components are used

Some programs are assessed in terms of qualitative features indicating that, the basic features remain unchanged.

There is a correlation between the quality features and re-configure action of the implementation program; while, the qualitative characteristics are determined by machine and cannot be determined in advance.

III MATERIAL

There exist Methods to calculate the impact of components where, if a client-server multimedia application is of concern the insertion of a component that compresses the data stream may optimize the end-to-end response time in case of limited communication bandwidth between client and server.

This phenomenon is not implemented as a result of a configuration program yet.

Within this newly introduced framework, adaptation policies defines a mapping of monitored parameters in order to apply configurations

The tool can be applied to define the profile adaptation.

The graphical tool is applied to re-configure the program.

This tool uses a re-configuration interface for each component and finally developments a package.

IV MODEL OF COMPUTATION

Dynamic reconfiguration framework is based on an algorithm first introduced by Kramer and Magee and Wermelinger optimized this algorithm [8, 13].

The model applications in this article consist of interconnected computational entities, named: in-port and out-port. The types of components are active and passive. The maps can be plotted by directed graph components, where nodes represent components and arcs represent the connections; this map should be Acyclic.

Every component is connected to at least one component. The state of the component is changed through a contact with other components. In order to preserve global application consistency during a reconfiguration, the communication among components must be tracked the concept of component transaction is introduced in this article.

A transaction consists of a series of two-way conversation between components which should be completed in a limited time. The initiator must be notified about the end of the transaction.

A transaction is related to other transactions, if completion is related to the implementation of other transactions.

Four connected components are shown in Fig (2).

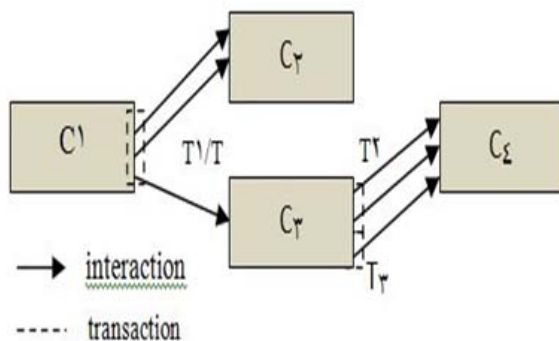


Figure2. Application model [3]

Transaction, T1 depends on transaction T2, illustrated as T1 / T2, Fig (2). Transaction T1 is not complete until transaction T2 is completed. Assume integrated client-server and proxy in Fig (2) where, the interactions between the client and proxy depend on interactions between the proxy and the server. The initial transaction is completed only when the transaction is completed between the Proxy and Server. Each arrow represents a transaction between the components. At the implementation level flexibility is possible.

A variety of connectors ranging from simple shared memory and TCP/IP sockets to .NET Remitting and CORBA connectors, are implemented, subject to exchange during runtime through dynamic reconfiguration. The interaction between the components is takes place by connectors.

V ALGORITHM FOR DYNAMIC RECONFIGURATION

Component-based dynamic re-configuration has a number of unique operators like removing component and modification of component parameters.

More complex operations like updating components and migration to other hosts on the runtime components should be kept in consistency in their state in order to continue the program after the re-configuration some levels of consistency are desirable in detail in Ensuring dynamic reconfiguration consistency [15].

All ports must be connected and the affected components must have adaptation input or output connections. The structural integrity should be evaluated by statistical tools like UML In the following an extension of the deadlock prevention algorithm is introduced in order to enhance and ensure the isolation property in case of reconfiguration.

Local and global adaptation is guaranteed by the following algorithm at runtime:

```

if (Ti before Tj)
If (Tj == reconf) Ti is rollbacked
else if (Ti == reconf) Tj is rollbacked
else Ti wait
else if (Ti = reconf and Tj!=reconf)
Tj is rollbacked
else Ti is rollbacked
    
```

This algorithm is based on the fact that when a transaction t is conflicting with a reconfiguration transaction tr, t is roll backed despite the order between t and tr. This algorithm ensures that application transactions remain consistent despite a concurrent reconfiguration transaction.

Kramer and Magee in their first algorithm described a process named "freezing" where you the activity of all components are stopped [9]. Wermelinger improved the "freezing" process by blocking only the connections subject to reconfiguration [13]. One of the advantages of this process is in its reducing the block time in all connections. Wermelinger explained that there exists the possibility of re-configuration even if all the connections in transactions in progress are blocked. It is expected

from a block transaction to continue its progress after blocking is finalized.

In order to ensure that all transactions are blocked the dependent blocks must be blocked as well otherwise, the deadlocking occurs. A new blocked component can be re-connected to the components if it is added or changed to fit the component parameters.

In order to reduce the block time number of steps of reconfiguration, loading before blocking and removal of the component after blocking should take place.

VI RECONFIGURATION OF REAL-TIME APPLICATIONS

As mentioned, consistency of state should remain unchanged for dynamic reconfiguration. This newly introduced algorithm guaranties this claim.

Provided that the CPU has a time schedule

Orders must be re-configured in a limited time

In general, the real-time applications, is made through periodic time [5].

In order to dynamically re-configure command processor resources to be reserved in a dynamically manner resources should be calculated with respected to special rules in order to reduce the overhead

The reconfiguration, as usual, including the activities of non-real time

The blocking is limited to reconfiguration activities like loading the new components or removing the old components which may need unknown running time; hence, they must be performed before or after the time block.

It is assumed that this strategy may be final, that is, the conflict between old and new components and there is always exists a trade-off between the block and re-configuration time.

In the first stage the details must be analyzed. Re-configuration of the system can be calculated through Eq. (1) [3].

tr: time between the reconfiguration request and the new configuration to run

tb: blackout time - time the application is interrupted

tl: time to load and initialize new components

td: time to delete components not contained in new configuration

$$Tr = tl + tb + td \quad (1)$$

When loading new components their details must be re-configured.

To support real-time properties in component-based systems (including not only their execution but also their replacement), components are analyzed offline to calculate their loading, unloading, and replacement times[7].

To calculate the reconfiguration time the blackout time must be considered in detail with no data processing in progress.

That this time should be considered when designing is of essence.

Blackout time consists of the following sub-time periods:

Te: The time required for reconfiguration commands

Ti: The maximum time to complete the transaction i

n: The number of transactions in the application

Tstate: State transfer

Tci: Time initialize and start connecting

Ts: Time to restart the application

Blackout time is calculated as follows:

First, do not count on state transfer

Note that the time remaining for transaction execution with blackout time is limited.

Runtime reconfiguration depends on the number of components, of attributes and number of connections. Each of these commands contain a limited number of variables that can be implemented in a limited time

It should be mentioned that the delay or cache memory is not considered in this study.

The calculation of the number of components involved in blocking and the connections / transactions subject to blockage must be conducted beforehand, as well as the timing instructions for reconfiguration.

If all software settings detected before time runs, the entire schedule changes can occur offline

Block connections of all transactions are in progress through the connection. When all transactions are executed in parallel block the transaction is limited to the time of the transaction, represented by $\max(t_1 \dots T_n)$.

This time depends on the transactions involved in interactions.

Upper bound for dialog transactions must be analyzed before block out time runs out. If developers adhere to certain rules of application, the downtime will be minimal. The reconfiguration method proposed here for small programs.

To develop new components, reassignment of connection and connectors are required; moreover, the CPU resources must be re-configured due to limitation of running time [11]. Here, synchronization is necessary between a component subject to reconfiguration and its connection.

The possible manners to change the task by task configuration are detailed in [10].

The reconfiguration approach Tasks is components and connection is the synchronization between them

The advantage of this newly proposed algorithm is that it is independent of the actual implementation. Implementation of any type of connection at the time of completion and their reconfiguration during runtime is made possible here.

The real-time control software is illustration in Fig. 3. A component processes the external sensor data and sends the result to the next component, which should generate the output data at every 20 milliseconds.

The third component, insertion is illustrated in Fig 4. This component generates logging data for debugging purposes.

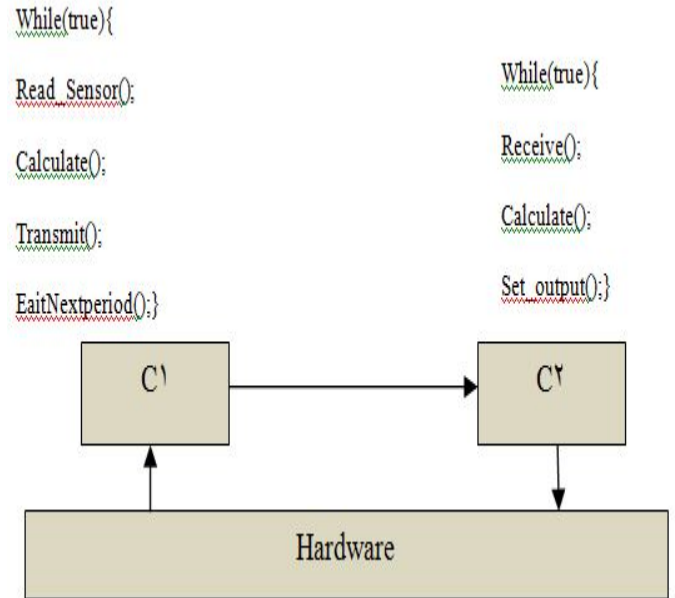


Figure3. Simple Real-Time Application [3]

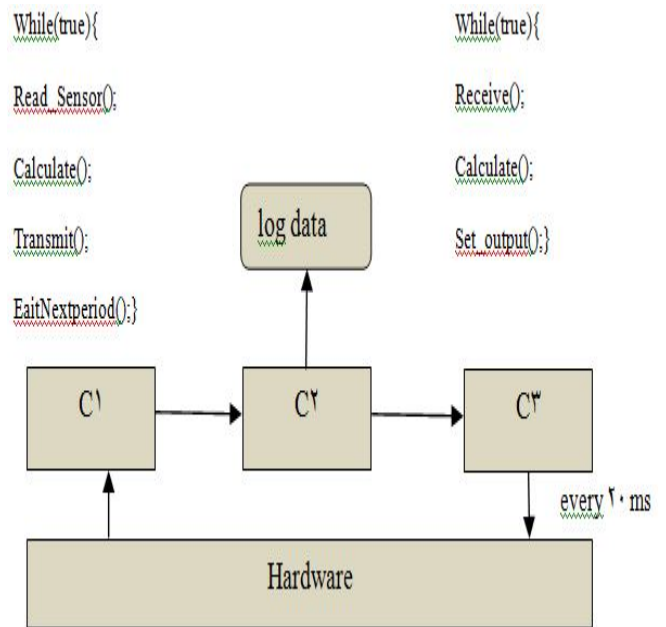


Figure4. Added Debug Component

The application is re-configured in a dynamic manner. Assuming the worst case of 5 ms for the first component and 5 ms for the second component the 10 milliseconds need to be re-configured.

Of course, the logging component is not allowed to use more than 5.6 ms if a rate monotonic scheduler is of concern, which requests an upper bound for CPU utilization.

VII RECONFIGURATION AND STATE ITS AFTERMATH

In case of migration or transition updating of the component should be considered.

To maintain the compatibility mode the state transfer should be considered.

Transfer mode must be applied in the blackout time; otherwise, the re-configuration will be impaired.

When re-configuration for analysis are considered the transfer time, there exist several methods in handling state transfer in presented scenarios.

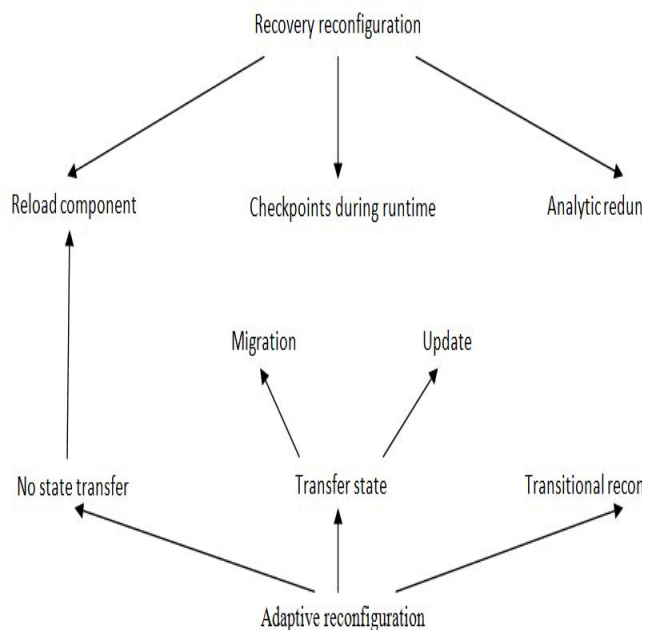


Figure5. Reconfiguration and State[3]

An overview of the implementation of a state during the re-configuration is shown in Fig 5. When all the components in an error state of the component disappear, a reconfiguration could be triggered to reestablish a valid configuration of the application.

Software is applied for real-time read data from sensor input, process and generate make the output signal. In this process, if the emphasis is not on the state component of the crash of component re-load the same component. One method of dealing with the crashed components stored them periodically. These checkpoints are used to set up new components in case of error. The disadvantage of this method is its having a high overhead and lack of overall consistency. One of the best methods to implement components of the error is using the redundancy analysis application.

In the analytic redundancy, different algorithms are available in generating different components.

If one of the redundant components defaults, it is replaced by a backup component, leading to generating a correct output.

Such a state needs no transfer because redundant components are controlled in parallel.

In this study this method is applied in the laboratory. If reconfiguration is requested by an adaptation policy due to changes in the environment, the state of component instances is still available. If the reconfiguration requires components to be migrated or updated, state of old component instances must be transferred to new ones.

If a re-configuration is of concern environmental compatibility will be available as a component state; while, if a re-configuration is required it makes the migration of the old state component necessary.

In our existing implementation, state transfer was realized by using a reflection interface for component instances and simply copying values of all instance members recursively to migrated and updated instances respectively. In case of new members in new component versions, these fields are initialized with default values. Missing members are not considered, any more and their state is left over.

The reflection of the values interface is applied and the copies of the components are implemented.

The time required to move a component to another component depends on the size of both the components. E. Schneider et al have adopted this method [6].

During the re-configuration state component is transmitted to the update service component.

If there are requests that must be handled before state transfer completes, these requests will be processed first and the state transfer is restarted afterwards. This approach works fine if work load in single periods varies significantly. Otherwise the maximum instance size of a component must be specified by component programmers. State size is always bounded by available memory in the target system. It is assumed that there exists enough memory to load the factors and the transition state. This method works well when a specific period is assigned to the workload. For a limited time both the new and old components run in parallel.

In general, some state component does not require state transition. The new component is initialized in a block-out. When the new configuration is run, the old components are deleted.

VIII RELATED WORK:

The original work by M. Wermelinger has been on a theoretical basis and was not targeted at real-time control software. Initially we focused on the implementation of Wermelinger's algorithm, which later on has been adapted to our problem specification.

The Swift toolkit enables dynamic reconfiguration of feedback control applications by limiting component interaction to a simple input/output model [1]. Reconfiguration is performed by changing component parameters. Adaptive behavior is reached by guarding system parameters.

E. Scheider et.al describe the dynamic reconfiguration real-time service-based software through the OSA+ middleware [6]. With OSA+, service reconfiguration is scheduled as a normal task. This allows the specification of deadlines and priorities for reconfiguration requests. During reconfiguration of a service, the old services' state is transferred to a newly loaded instance. If any service requests have to be scheduled during this phase, because of a close deadlines state will get inconsistent. That is why the reconfiguration task restarts the state-transfer to preserve service consistency. If the state has been transferred without interruption, the new service instance will be activated - new service request then are forwarded to the new service instance.

This is a very powerful and general approach for service-based architectures; however, the reconfiguration of services is very coarse grain in

contrast to a reconfiguration based on application objects, as presented in this paper.

M. Pfeffer and T. Ungerer present an approach for dynamic reconfiguration on a multithreaded Javamicrocontroller, Komodo [12]. The Komodo microcontroller is able to execute Java byte code directly and run multiple threads parallel in hardware. Single Java objects can be updated using a separate thread. Updates are realized by overwriting an object's code memory section and updating method invocation tables and object data if required. The authors give a worst case switching time of 246 clock cycles. With their approach, it is only possible to update one object at a time. The Simplex approach introduced by Lui Sha et.al [14, 10]. describes a special case for dynamic reconfiguration, which has already found its way into highly responsive systems like the Boeing 777 flight controller. The idea is to use analytic redundancy for control. With Simple, faults in a high performance controller trigger a dynamic reconfiguration

That activates a more reliable controller instance implementing a simpler control algorithm, but still producing valid output signals.

Mode Changes for priority-driven scheduling algorithms describe task set reconfiguration for real-time systems [13].

Tasks can be added, modified or removed from an existing task set. Mode change protocols define rules when reconfiguration can take place and how, including the adjustment of task synchronization (e.g. ceiling value adaptation in the priority ceiling protocol). The reconfiguration from one task set to another is split into single operations (add, modify, remove), which are executed sequentially. A reconfiguration can last several task periods. Our approach is equivalent to mode changes if we consider tasks as concrete component instances and synchronization as their interaction. Our approach is more general.

IX CONCLUSIONS AND FUTURE WORK

Within this paper, we have presented our framework for dynamic reconfiguration that allows applications to adapt to changing environmental conditions. Based on prior work we have here described the steps necessary for reconfiguration. Furthermore, we have analyzed their timing behavior, which can be calculated for many application configurations.

Using our approach it is possible to reconfigure complex component-based applications with complex component

dependencies. Dynamic reconfiguration is not limited to the simple exchange of a single component, but instead allows for dynamic modification to complex parts of an application. Reconfiguration and state transfer have been discussed for different scenarios. We have analyzed the impact of these procedures on the time bounds of black-out periods.

We have demonstrated that the blackout time during reconfiguration is bounded and can be calculated for given applications. This allows for the construction of real-time software that is reconfigurable dynamically, while all task's deadlines are met. It is possible to show that the used reconfiguration algorithm a generalization of the mode change theory, which has been theoretically proven.

The presented adaptation strategy has been integrated into an experiment controller within the Distributed Control Lab, a remote laboratory, to ensure safety against malicious code downloaded from the Internet.

We have implemented and evaluated the described approach based on standard middleware components. We are now going to implement a new version of our framework for real-time control software. We plan to report on measurements in a separate publication and will describe our experiences made in context of our remote lab.

REFERENCES

- [1] A. Goel, D. Steere, C. Pu, and J., "Walpole A feedback control and dynamic reconfiguration toolkit," In Proceedings of the 2nd Usenix Windows NT Symposium, pages 172–172, .1998.
- [2] A. Rasche and A. Polze, "Configurable Services for mobile Users. In Proceedings of IEEE Workshop on Object Oriented Real-time Dependable Systems," pages 163–171, San Diego, CA, 2002.
- [3] A. Rasche and A. Polze, "Configuration and Dynamic Reconfiguration of Component-based Applications with Microsoft .NET," In International Symposium on Object-oriented Real-time distributed Computing (ISORC), pages 164–171, Hakodate, Japan, 2003.
- [4] A. Rasche, P. Tröger, M. Dirska, and A. Polze, "Foucault's Pendulum in the Distributed Control Lab" In Proceedings of IEEE Workshop on Object-Oriented Real-time Dependable Systems, pages 299–306, Capri Island, Italy, 2003.
- [5] C. Liu and J. Layland, "Scheduling algorithms for multi programming in a hard-real-time environment" Journal of the ACM, 20(1):46–61, 1973.
- [6] E. Schneider, F. Picioraga, and U. Brinkschulte, "Dynamic Reconfiguration through OSA+, a Real-Time Middleware" In Middleware 04, 1st Middleware Doctoral Symposium, Toronto, 2004.
- [7] J.C. Romero, M. García-Valls, (2013) "Scheduling component replacement for timely execution in dynamic systems" (16 December 2012) [online]. Available: wileyonlinelibrary.com
- [8] J. Magee, N. Dulay and J. Kramer, "Specifying Distributed Software Architectures," In W. Schafer and P. Botella, editors, Proc". 5th European Software Engineering Conf. (ESEC 95), volume 989, , 2012, pages 137–153, Sitges, Spain., Springer-Verlag, Berlin.
- [9] J. Kramer and J. Magee, "The evolving philosophers problem: Dynamic change management", IEEE Transactions on Software Engineering, 16(11):1293–1306. 1990
- [10] L. Sha, "Using simplicity to control complexity", IEEE Software, 2001, 18(4):20–28.
- [11] L. Sha, R. Rajkumar, J. Lehoczky, and K. Ramamritham, "Mode change protocols for priority-driven preemptive scheduling. Real-Time Systems", 1989.
- [12] M. Pfeffer and T. Ungerer, "Dynamic real-time reconfiguration on a multithreaded java-microcontroller". In International Symposium on Object-oriented Real-time distributed Computing (ISORC), 2004 pages 86–92, Vienna, Austria. IEEE.
- [13] M. Wermelinger, "A hierarchic architecture model for dynamic reconfiguration", In Proceedings of the Second International Workshop on Software Engineering for Parallel and Distributed Systems, 1997 pages 243–254, 1109 Spring Street, Suite 300, Silver Spring, MD 20910, USA, IEEE.
- [14] M. Bodson, J. Lehoczky, R. Rajkumar, L. Sha, M. Smith, D. Soh, and J. Stephan, "Control reconfiguration in the presence of software failures", In Proceedings of the 32nd IEEE Conference on Decision and Control, volume 3, 2012, pages 2284–2289.
- [15] N. D. Palma, P. Laumay, and L. Bellissard, "Ensuring dynamic reconfiguration consistency", In 6th International Workshop on Component-Oriented Programming (WCOP 2001), ECOOP related Workshop, 2001 pages 18–24, Budapest, Hungary.
- [16] IIOP.NET Homepage. .NET, CORBA and J2EE Interoperation. [online]. Available: <http://iiop-net.sourceforge.net>, 2004. [Accessed: October. 12, 2015].
- [17] w.li, QoS Assurance for Dynamic Reconfiguration of Component-Based Software Systems, VOL. 38, NO. 3, IEEE Computer Society, 2012

Intrusion Detection Systems: A Novel Multi-Agent Specification Method

Eddahmani Said

University Mohammed V Rabat

Faculté des Sciences de Rabat, Morocco

Rahal ROMADI

University Mohammed V Rabat

ENSIAS, BP 713, Agdal Rabat, Morocco.

Bouchaib BOUNABAT

University Mohammed V Rabat

ENSIAS, BP 713, Agdal Rabat

Abstract

Intrusion detection systems play a major role in information systems security. However, the complexity of a network flow can cause the outbreak of numerous false alarms, commonly called false positives or cannot detect malicious traffic, i.e. false negatives. Our objective in the present paper is to put forward a novel method that combines several types of monitoring, intrusion and malware detection tools whose main role is to reduce the rate of false positives/negatives.

Keywords: *malware detection system, Reactive Agent, specification, formal methods.*

I. INTRODUCTION

The Internet has become indispensable in modern society, where everyone can receive the services provided such as access to services and information (cloud) despite geographic barriers. However, besides the benefits, the users can become potential victims of various types of cyber threats. Intrusion detection systems (IDSes), intrusion prevention systems (IPsec), access control, firewalls, and honeypots are the security mechanisms implemented to secure networks or systems. But attackers are able to manage even these security mechanisms and trying to break into the legitimate system or network with the help of various evasion techniques. In this paper, we seek to exploit the complementarities between these tools. For this there is provided a centralized monitoring system wherein cooperates all types of malware detection system.

In this paper, we described a method of specification and detection system design. We consider that such a system is a reactive multi-agent system consisting of concurrent reactive agents that cooperate with each other to achieve the desired functionality. In

addition, this approach uses formal synchronous specification and verification tools in order to specify and to verify the systems behaviors.

Types of intrusion detection system

Basically we can limit the intrusion detection tools to the following types:

- **Signature Intrusion Detection System:** The simplest form of signature recognition uses simple pattern matching to compare the network packets against binary signatures of known attacks. A binary signature may be defined for a specific portion of the packet, such as the TCP flags. Signature-based schemes provide very good detection results for specified, well-known attacks. However, they are not capable of detecting new, unfamiliar intrusions. New virus attacks such as ADMutate and Nimda create the need for multiple signatures for a single attack. Changing a single bit in some attack strings can invalidate a signature and create the need for an entirely new signature. Despite problems with signature-based intrusion detection, such systems are popular and work well when configured correctly and monitored closely.

- **Anomaly Intrusion Detection System:** Anomaly detection differs from the signature recognition model. The model consists of a database of anomalies. Any event that is identified with the database is considered an anomaly. Any deviation from normal use is labelled an attack. Creating a model of normal use is the most difficult task in creating an anomaly detector. In the traditional method of anomaly detection, important data is kept for checking variations in network traffic for the model. However, in reality, there is less variation in network traffic and too many statistical variations making these models imprecise; some events labelled as anomalies might only be irregularities in network usage. The main benefit of anomaly-based detection techniques is their potential to detect previously unseen intrusion events. However the rate of false positives (events erroneously classified as attacks) is usually higher than in signature based ones.
- **Log File Monitoring:** A Log File Monitor (LFM) monitors log files created by network services. The LFM searches through the logs and identifies malicious events. In a similar manner to IDS, these systems look for patterns in the log files that suggest an intrusion. A typical example would be parsers for HTTP server log files that look for intruders who try well-known security holes, such as the "phf" attack. An example is swatch. These mechanisms are typically programs that parse log files after an event has already occurred, such as failed log in attempts.
- **File Integrity Checking:** These mechanisms check for Trojan horses, or files that have otherwise been modified, indicating an intruder has already been there, for example, Tripwire.
- **System Integrity Verifiers:** A System Integrity Verifier Agent monitors system files to determine whether an intruder has changed the files. An integrity monitor watches key system objects for changes. For example, a basic integrity monitor uses system files, or registry keys, to track changes by an intruder. Although they have limited functionality, integrity monitors can add an additional layer of protection to other forms of intrusion detection.
- **Firewalls:** The firewall verifies the incoming and outgoing traffic against firewall rules. It acts as a router to move data between networks. Firewalls manage access of private networks to host applications. All the attempts to log in to the network are identified for auditing. Unauthorized attempts can be identified by embedding an alarm that is triggered when an unauthorized user attempts to login. Firewalls can filter packets based on address and types of traffic. They identify the source, destination addresses, and port numbers while address filtering, and they identify types of network traffic when protocol filtering. Firewalls can identify the state and attributes of the data packets. Firewalls are mainly categorized into four types:
 - ✓ Packet filters: Traffic allowed based on source and destination IP address, packet type, and port number
 - ✓ Circuit-level gateways: Traffic allowed based on session rules, such as when a session is initiated by a recognized computer
 - ✓ Proxy/Application-level gateways: Traffic allowed based on specified applications (such as a browser) or a protocol, such as FTP, or combinations
 - ✓ Stateful multilayer inspection firewalls: Traffic is filtered at three layers based on a wide range of the specified application, session, and packet filtering rules
- **Honeypot:** A honeypot is a system that is intended to attract and trap people who try unauthorized or illicit utilization of the host system by simulating vulnerable system services and Trojans. By acting as a decoy server, it can divert attacks from critical systems and provide a higher level of information than can be achieved by using firewalls and NIDS alone. We can monitor attackers' tricks and exploits by logging their every activity, so that we can respond to these exploits quickly without attackers actually misusing and compromising systems.

- **Antispyware/antikeylogger:**

The presence of keylogger and spyware is usually synonymous with attack networks. One can find in the gifts of these tools very utile information to detect the origin of the attack.

The method we propose provides a way to specify a detection system based on these different tools. We consider that this system can be modeled as a distributed computing system consisting of several autonomous agents. Each agent represents a detection tool and these interactions with its environment. In this paper, the agents are classed as either deliberative or reactive. Deliberative agents derive from the deliberative thinking paradigm: the agents possess an internal symbolic, reasoning model and they engage in planning and negotiation in order to achieve coordination with other agents. Reactive agents don't have any internal symbolic models of their environment, and they act using a stimulus/response type of behavior by responding to the present state of the environment in which they are embedded.

The use of agent-orientation in the modelling, design, and implementation provides at least the following benefits:

- **Flexible.** Agent architectures are more flexible, modular and robust than, for example, object-oriented ones. They tend to be open and dynamic as their components can be added, modified or removed at any time.
- **Pro-activeness.** Intelligent agents are able to exhibit goal-directed behavior by taking the initiative in order to satisfy their design objectives:
- **Reactivity.** Agents are crucial when operating in an unpredictable environment containing a large number of data sources scattered over multiples sources. If an agent queries an information source and finds no answers to its query, it would then try alternate sources of information until it could come up with a reasonable number of answers.
- **Learning.** Another important characteristic of autonomous behavior is the ability to enhance future performance as a result of past experiences. Machine learning techniques allow an agent to learn new methods or refine existing ones to meet specific needs.

- **Communication and cooperation.**

Intelligent agents are capable of interacting with other agents (and humans) in order to o achieve a common goal.

II. SPECIFICATION AND VERIFICATION TOOLS

This section will describe all the specification and verification tools used in this work.

A. SynCharts

SynCharts (SC) are introduced by Harel [6] like a visual formalism that provides a way to represent state diagrams with notions like hierarchy, concurrency, broadcast communication and temporized state. A SC can be seen like one or several automata which are labeled by Event [condition]/ action. SC is said to be synchronous because the system reacts to events by instantly updating its internal state and producing actions, the actions produced can trigger in the same instant other transitions, this is named chain reaction causing a set of transitions, the system is always in a waiting state until the condition for a transition is true.

B. ESTEREL

To hit this target, the specified SC behaviors are automatically translated to the synchronous language ESTEREL [1][2][3]. It's a language, with precisely defined mathematical semantics, for programming the class of input-driven deterministic systems. The software environment of ESTEREL provides high-quality tools, including an editor, compiler, simulator, debugger and verifier.

C. Real-Time Temporal Logic

Temporal logic has been widely used for the specification and verification of concurrent systems [4][5]. However, these temporal logics only allow qualitative reasoning about time. Several extensions have been proposed for expressing and reasoning about real-time systems. These include Real-Time Temporal Logic (RTTL) [9], which is based on linear time temporal logic, and allows in addition the expression of quantitative real-time properties (e.g. exact delays or event deadlines).

Example of RTTL Formula

$s1 \wedge t = T \rightarrow \diamond (s2 \wedge t \leq T + 5)$ - If s1 is true now and the clock reads T ticks, then within T + 5 clock ticks, s2 must become true. Thus,

once s_1 becomes true, s_2 must become true no more than 5 ticks later. This formula can be also written as follows: $s_1 \rightarrow \diamond[0,5] s_2$ or $s_1 \rightarrow \diamond \leq 5 s_2$

The formula $s_1 \leftrightarrow s_3$ indicates that events s_1 , s_3 are simultaneous. If $C(w)$ is a RTTL formula defining a temporal constraint on an event w , then $w \models C(w)$ means that w satisfies the formula $C(w)$.

III. SYSTEM ARCHITECTURE

The internal organization of our system consists in a tree (fig. 1) that is made up in parallel of a supervisor (*Supervisory Agent*), of two or several sub-agents components, and two communication interfaces between the supervisor and the sub-agents.

Such system interacts with its environment by the means of:

- Actions exerted by this environment.
- External States emitted to the environment.

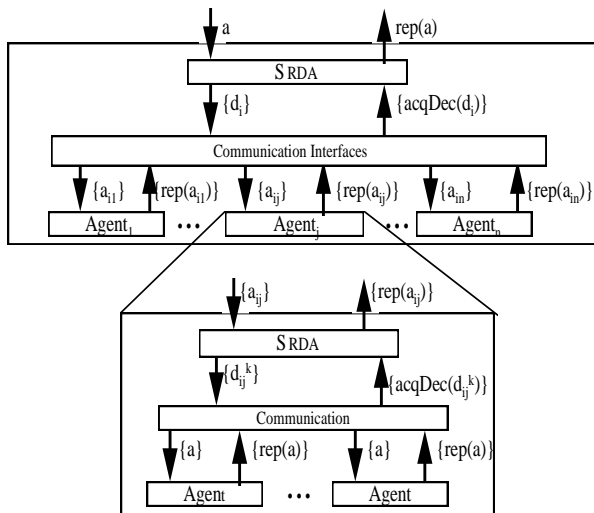


Fig.1. Architecture of our system.

A. Supervisory and Sub-Agents Levels.

The supervisory agent (SRDA: Supervisory Reactive Decisional Agent) is a RDA controlling the component sub-agents, in order to achieve a goal or to solve a given problem.

This agent will manage the sequences of activation and the definition of the controlled sub-agents objectives. This management depends on:

- The actions exerted by the environment,
- The events generated by the sub-agents activities,
- The temporal constraints specific to any reactive system.

In addition, a reactive system can be summarized with a simple SRDA directly connected to the controlled process. Each sub-agent can be considered as a reactive system. Thus, its internal structure is composed by its own SRDA, communication interfaces and sub-agents. Sub-agent objectives are to carry out sequences of tasks in response to any action exerted on him by the higher level.

B. Communication Interfaces

The communication interfaces are of two types: decisional interface that translates a decision (d) generated by the SRDA into several actions (a_i), each one of them is intended for a sub-agent of the lower level. Signaling interface that synchronizes the external states (e'_i), sent by each sub-agent, and emits one signaling (s) intended for the SRDA.

C. Formal Description

The proposed model of reactive agent consists in putting forward decisional models allowing the representation of objects according to their behavioral aspects and their degree of intelligence.

Definitions. A Reactive Decisional Agent (RDA) is 9-tuple noted $\langle A, D, S, E', O, O', act, dec, sig \rangle$ where:

- **A**: Set of actions exerted on the agent. Each action, undergone by an object, represents a possible operation to be carried out on this object in order to achieve a specific goal.
 - **D**: Set of decisions generated by the agent. Each decision is a solution concerning process behavior in the future; each decision is characterized by its action horizon: H_a , the time during which this decision remains valid.
 - **S**: Set of Signaling received by the agent. Each Signaling received by an object, reflects at any given time the state of the controlled tools used to achieve a specific goal.
 - **E'**: Set of external states delivered by the agent. Each one represents the object state emitted to the environment.
 - **E**: Set of agent's internal states. Each one indicates the current state of the agent.
 - **O**: Set of agent's internal objectives. Each decision is elaborated in order to achieve an internal objective according to the current external objective and the actual internal state.
 - **O'**: Set of agent's external objectives that can be achieved. These objectives represent the agent's interpreting of each action.
- From a dynamic point of view, the sets above indicate the received events (A, S), the emitted

events (D, E') and the internal events (E, O, O').

Decisional Functions. *act*, *dec*, and *sig* are three decisional functions that define the behavior of a RDA.

$$\begin{aligned}
 act: & A \longrightarrow O' \\
 a & \longrightarrow o' \text{ with, } \forall a \in A, \exists ! o' \in O' / \\
 o' = act(a) & \Rightarrow a \leftrightarrow o' \quad (1)
 \end{aligned}$$

(1) means that the occurrence of an action *a* implies instantaneously the occurrence of its associated external objective *o'* by the function *act*.

$$\begin{aligned}
 dec: & O' \times E \longrightarrow D \times O \\
 (o', e) & \longrightarrow (d, o) \text{ with, } dec(o', e) = (d, o) \\
 \Rightarrow & [o' \wedge e \leftrightarrow d \wedge o] \quad (2)
 \end{aligned}$$

(2) means that depending of the current external objective *o'* and as soon as the agent is in an appropriate internal state *e*, corresponding decision *d* an internal objective *o*, by the function *dec*, are instantaneously produced.

$$\begin{aligned}
 sig: & O' \times O \times S \longrightarrow E \times E' \\
 (o', o, s) & \longrightarrow (e, e') \text{ with, } sig(o', o, s) = \\
 (e, e') & \Rightarrow [o' \wedge o \wedge s \leftrightarrow e \wedge e'] \quad (3)
 \end{aligned}$$

(3) Means that depending of the current external objective *o'* and the expected internal objective *o*, and as soon as the receipt of a signaling *s*, its associated external state *e'* is instantaneously emitted and the new agent internal state becomes *e*.

According to the formal definitions above, figure.2. shows the internal structure of a Reactive Decisional Agent. Act interprets an action as an external objective, that it used by Dec and Sig to generate agent appropriate responses.

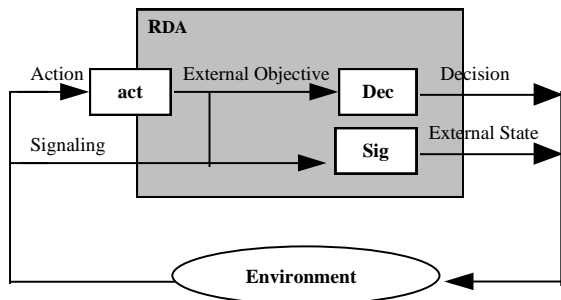


Fig.2. Internal Architecture of an RDA

This section presents a set of SC which describe the external objective of a RDA.

External Objectives Manager. A Reactive Decisional Agent has an External Objective Manager. It consists in a SynCharts model of the function act described above (Fig. 3).

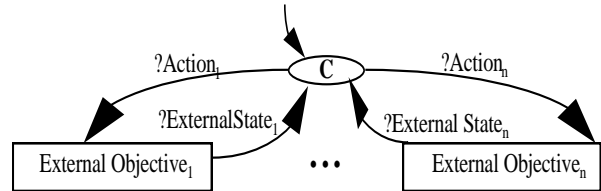


Fig. 3. Syncharts model of External objectives manager

In the Syncharts model of External objectives manager, each state represents an external objective whose activation is started by the reception of a specific action (*?Action*), and terminated by the emission of the acknowledgment external state (*!ExternalObjective*).

External Objectives Modelling. An external objective is composed by many others SC states corresponding to the associated internal states and internal objectives that are deduced by the functions *dec* and *sig* definitions (Fig. 4).

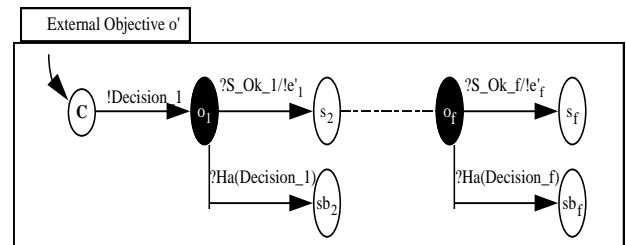


Fig.4. The general SC model of an External objective

The Figure (fig.4) shows the general SC model of an External objective. The transition (*Internal state* → *Internal objective*) is made by a decision emission (*!Decision*), and the transition (*internal objective* → *Internal state*) is made by a signaling receipt (*?S_OK*), and eventually an external state emission (*!e'*). Internal state C corresponds to the default initial state of a SC model. Internal state and internal objective are indicated respectively by *e_i* et *o_i*. In case of an action horizon exceeding without receiving any acknowledgment signaling, the agent's internal state changes from *e_i* to *eb_i* (*breakdown state*).

IV. Conclusion

The contribution of this paper is to give a new formal approach to specify and design a malware system detection. The originality is to consider each existing detection tool as a Reactive Decisional Agent, and to bring together several formal synchronous modeling and validation tools. With its top-down process and its principles of decomposition, this method allows to get a model which is more easily understandable by the user. The SynCharts models are used here in order to describe the reactive agent behaviors. These behaviors will be checked in a qualitative (respectively quantitative) way by the synchronous language ESTEREL (respectively by Real Time Temporal Logic deduction).

REFERENCES

- [1] Berry, G.: The ESTEREL V5 Language Primer. Internal Report, CMA Ecoles des Mine, INRIA, Paris, 17 Mars (1998).
- [2] F. Boussinot, and R. de Simone. : The ESTEREL language. Proceeding of the IEEE, 79(9):12931304.
- [3] Berry, G. and P. Couronne. : Synchronous programming of reactive systems: an introduction to ESTEREL. IEEE Software Engineering.
- [4] Pnuelli, A.: The Temporal Semantics of Concurrent Programs. Theoretical Computer Science, 13 pp. 4560 (1981).
- [5] Ostroff, J.S.: Temporal Logic for RealTime systems. (Advanced Software Development Series. Research) 1st edn, (1989).
- [6] D. Harel, M, Politi, Modeling Reactive Systems with Statecharts: The STATEMATE Approach, McGraw-Hill, ISBN 0-07-026205-5.
- [7] Guide to Intrusion Detection and Prevention Systems (IDPS), NIST special publication 800-94, by Karen Scarfone Peter Mell

A Brief Overview of Cooperative MAC Protocols in Wireless Networks

Ferzia Firdousi

Department of Electrical Engineering,
National University of Sciences and Technology
Rawalpindi, Pakistan

Abstract—Cooperative diversity is a fairly recent technique. It allows the benefits of space diversity to be incorporated into the wireless communications network. Whereas it has found widespread use in network, MAC and physical layers, this paper highlights its implementation in the MAC layer as a researchable area. A wide number of techniques have been deployed in the MAC layer with a scope for good cooperation for the modern wireless networks. While all of them have been found to be useful, a few of them have disadvantages as well. In this paper, we present a brief overview of the MAC protocols deployed in the cooperative diversity of wireless communication. This survey is compact enough to understand properly but detailed enough to provoke research ideas.

Keywords—Cooperative Diversity, Mac Protocol, Energy Consumption

I. INTRODUCTION

Hardware complexity in the network, mainly caused by antennas, wires and cables etc., limited advancement for communications companies [1]. The need for a mode of communication which was independent of hardware was then felt [2]. In this regard, proposition of wireless techniques offered a solution to the large-scale and expensive issue and hence spatial diversity techniques in wireless networks, surfaced [3], [4].

Moreover, to accommodate the modern wireless need for independent networks, distributed systems have to be considered. Such requirements are fulfilled both by ad-hoc networks or sensor network, both of which are sometimes called wireless distributed networks (WDN).

Ad hoc networks contain both mobile and immobile nodes. This means they cannot rely on an existing infrastructure and rather require a radio device to form at least a single hop network with peer radio devices to distribute and compute a task. So the assignment of each node, apart from transmitting its own packets is to forward packets for other nodes as well. This in other

words is called routing. Hence the next question concerns which nodes to be used for forwarding the data for these other nodes. This decision is made on the basis of network connectivity. The positive outcomes of the tasks include reduced per-node computational latency, energy and power consumption. This purpose of ad-hoc network requires a MAC protocol which not only schedules transmissions among the radio devices efficaciously but also governs data distribution among peer radio devices and data retrieval from them.

The technique to specify this forwarding property of networks is cooperative communication. One of its more popular types is cooperative diversity which is a type of space diversity that allows the use of virtual MIMO and cooperation techniques [5]. It can also be said that cooperative diversity decreases the need for hardware in a wireless system. The network, MAC and physical layers, all play a role in cooperative communications. This paper concentrates on MAC layer. It also looks into the trade-offs that researchers have to face when looking for the best solution for the problems faced while incorporating MAC protocols with cooperative communications in wireless networks [6].

Cooperative communication is easily implemented in ad-hoc networks because of its infrastructure-less nature. The broadcast and cooperative nature of ad-hoc networks are plus points for cooperation [7], [8]. By mitigating the two main causes of performance degradation in wireless transmissions; channel fading and signal interference, cooperative communications allows the distributed wireless networks to attain high spatial and multiplexed diversity gains [9].

With cooperative communications, networks work by allowing the idle nodes to relay the signals emitted by the source to the destination as shown in Fig. 1 where $H_{s,d}$, $H_{s,r}$ and $H_{r,d}$ are the impulse responses

of source to destination, source to relay node, and relay to destination nodes respectively [10]. Some relays may have a shorter distance destination than the source and some may have larger as shown in Fig. 2. Each of the relay paths are shown by dotted lines. It can be seen that a relay which may be closer to the source will not necessarily be close to destination for instance Relay 1 in Fig. 2. Therefore in this situation it is better to use the direct (Source to Destination) link. If the correct relay is used (or not used) the chances for a faded signal to reach the destination are greatly reduced [11].

With the inclusion of cooperation throughput is maximized [12], interference area is minimized [13], effective transmission rate is also increased [14], energy is conserved and the radio channel is effectively reused [15].

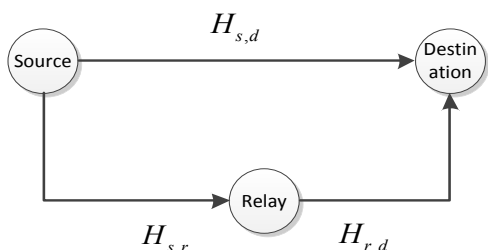


Figure 1. A cooperative diversity protocol layout

The most important points to note in a cooperative communication setup include optimal power allocation, relay selection and relay placement strategies [16]. All these points should be designed in a way to ultimately reduce the energy consumed by the system while at the same time increasing throughput rate.

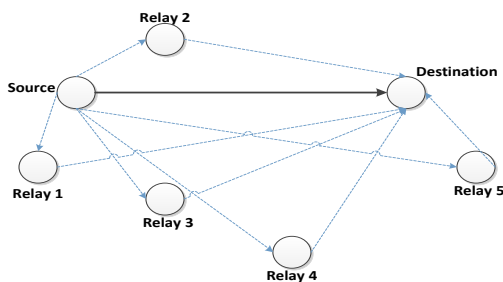


Figure 2. An array of relay nodes. Best relay is selected for cooperation to reduce fading.

MAC protocols in ad-hoc networks generally devised in previous research are classified into two groups; a) MAC protocols which use an indirect link (via relay) when that link is more suitable for transmission than direct link, and b) MAC protocols which only use an indirect link (via relay) when the transmission via direct link fails. In the first case, the

relay node is pre-selected before transmission whereas in the second scenario the relay node is selected after an initial failed transmission via direct link. If one of these options is explored than the protocol is more suitably called cooperative relaying? If diversity has to be increased then doing either of these steps is not enough as the MAC protocol in either condition doesn't reap the benefit of the diversity gain [17]. A very good research path can be found by trying to incorporate both direct and indirect links in cooperation and using diversity combining techniques to achieve a good transmitted signal at the destination. This way both cooperative relaying and cooperative diversity are used together and benefits of both are reaped.

II. IMPLEMENTATION OF MAC PROTOCOL: WHY IS IT DIFFICULT?

While implementing MAC protocols a lot of things need to be kept in mind including energy efficiency, its ability to be incorporated into a standardized protocol and adaptability to different international standards [18]. If even one of these is missed, then implementation with cooperative communication becomes difficult [19]. When implemented with cooperation, MAC protocols tend to have energy consuming factors which waste energy and shorten network lifetime. These are explained below:

A. The hidden terminal

While a pair of transmitter and receiver (A and B) are communicating, and nodes such as C in Fig. 3 cause a collision at one of them because they do not know about the transmission at the pair. This way C is considered to be a hidden terminal.

B. The deaf terminal

As given above in the figure, again, to be a deaf terminal a node has to be unable to interpret and identify an ongoing transmission between a transmitter receiver pair such as A and B. so in this case the deaf node is C once again.

C. The exposed terminal

For a transmitter–receiver pair such as A and B, to be an exposed terminal a node will have to perceive signals sent by the transmitter, but not be able to cause a collision at the receiver for example when B sends a CTS signal it is received by both A and C. Node C in Fig. 4, is an exposed terminal.

D. Mobile Nodes

In the modern wireless technology, nodes (or station) are constantly moving, which gives rise to frequent disruption in connection from source/destination. This is also called node mobility. In

this case, relays selected in one channel frame do not remain the best relays in the next frame. This way judgment based on relay selection becomes redundant and has to be done all over again.



Figure 3. A hidden terminal

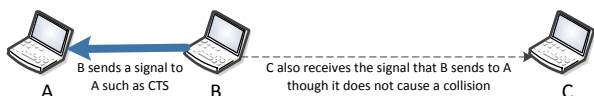


Figure 4. An exposed terminal

E. Idle listening

It occurs when nodes have to constantly listen to the channel for idleness in contention-based MAC protocols. This procedure wastes a large amount of energy on a single node which is waiting in queue to send its data. This node may have to listen for a very long time for the channel to get free for its transmission. But it is also a problem because the node has to listen to the channel irrespective of the fact that it has a packet to send or not.

F. Overhearing

It occurs when the nodes hear what is not intended for them. For example, in Fig. 3 the node C receives a signal forwarded from B to A. This packet is then wasted and so is the energy required to transmit it.

G. Protocol overhead

Protocol overhead is the name given to energy lost in transmission/reception of control frames. The control packets like RTS/CTS are designed to ensure that there is no idle listening or collision problem. This gives rise to a new problem which is also called packet overhead. It wastes energy.

H. Data packet size

The size of the data packet affects the receiving or emitting node and consumes more (or less) energy. The bigger the packet the more energy required for transmission/reception.

MAC protocols should be designed so that distributed nature of the network is maintained, real-time traffic is easily run; the access delay is minimized; bandwidth is efficiently utilized; each node gets a fair amount of bandwidth irrespective of hierarchy; protocol overhead is lessened; there are no hidden, deaf or exposed terminal problems; is scalable and compatible with existing technology; minimizes power

consumption; and provides synchronization between nodes. Therefore in order to have a reliable protocol which is fairly shared or available to each user, MAC is very essential for the network.

In the different MAC categories ranging from reservation-based and contention-based to fixed assignment and random-access [20], it is seen that what makes them different is the traffic, application, network-size or performance (SNR) requirements [21]. Each of these protocols faces different problems while being implemented. Altogether, there is no way that these methods can be perfectly implemented. But when compared with trade-offs reservation-based MAC protocol give better results. These are explained as follows:

A. Reservation-based:

In reservation based MAC, each transmission is scheduled and no nodes have to contend for the channel because each of them has their specific turns for transmission. Therefore only one transmitter-receiver pair communicates in each transmission period. This reduces collisions and eradicates idle-listening which conserves energy. But in order to maintain a specific order of transmissions, each node has to be perfectly synchronized with sensors and networks and this requires the maintenance of a large overhead. Usually different slots (based on time or frequency division) are used to separate different users. Therefore, the overall latency increases resulting in more energy wastage. But it is considerably better because it has less energy problems caused by contention.

B. Contention-based:

This protocol is easier to apply than reservation based schemes and also easy to synchronize. It usually involves Carrier Sense Multiple access. The main idea behind the scheme is that nodes keep listening for the availability of the channel at all times. This means that if any node finds the channel idle, it can contend it. But while this is true, other channels which are looking to transmit can also contend the channel. This would result in a collision. Thus the rate of collisions is high. This system is highly flexible and works best when the system is uncongested. But with a lot of traffic, it degrades rapidly and can cause energy loss.

III. IMPORTANT RESEARCH ISSUES OF COOPERATIVE MAC DIVERSITY PROTOCOLS

Here the question asked is “is cooperation necessary?” If the answer found is yes, then while integrating MAC protocol with cooperative diversity it should be noted that when selecting the best relay it is

necessary that complexity should be avoided and cooperation should only be done if it is beneficial [22]. Otherwise relay introduction is just another way of introducing more overhead and more interference. This would logically mean that all the benefits gained by the use of extra relays are eventually eradicated. In other words, it is necessary that relay nodes be used when they provide a better path for transmission, a better packet delivery, a faster throughput or less fading [23]. This means that relays which lie at an appropriate distance to both the source and destination are used for cooperation if required [24].

Having decided whether cooperation is necessary and beneficial, the next questions become 'how to cooperate' and 'whom to cooperate with'. The outcome required defines the methodology of selection of relay. For instance, if the requirement is to increase the effective transmission rate, channel conditions should be considered and thus a relay which can provide that is selected [25]; to maximize the network lifetime, using a relay with the most residual energy is preferred [26]; to improve the spatial reuse, the relay with the least neighbouring nodes is favoured [27]; to maximize the overall throughput, the relay that can achieve the highest cooperation gain and incur the smallest interference area should have the highest priority. While deciding which relay to use for cooperation it is necessary to check that the relay has a positive impact on the performance of the other nodes/sources/destination [28]. The relay chosen may be causing collision in the transmission of other sources, which would render its inclusion as unhealthy.

It is also commonly seen that relays are only used when needed [29]. Particularly, it should be noted that relays are used when, a) transmission via the direct link is not successful due to channel fading, packet loss or interference, or, b) when data packet transmission via indirect link (relay) gives better result direct link. This may be beneficial but it leaves the diversity advantage of relay in the back.

In a given wireless setup, there is no particular way of differentiating relay nodes from source nodes and this major issue is rarely mentioned in any of the research done so far. Moreover in these systems if a single relay is determined as the optimal one, how do the other (relay) nodes find out? This is particularly important in an ad-hoc network which is decentralized in nature. In TDMA-based MAC protocols with cooperation the question of when the relay should transmit also arises since all the time slots have been assigned to sources arises. While working with TDMA researchers often forget to mention this important point.

IV. CRITICAL ANALYSIS, TRADE-OFFS AND CONCLUSION

This paper analyses the scenarios governing the MAC protocols for cooperative diversity. There is a wide list of methodologies introduced for this purpose including CoopMAC [30], IreMAC [31] MS-C-TDMA [32], C-TDMA [33], ADC-MAC [34], VC-MAC [35], DRRS-MAC [36], CoopMC-MAC [37], C-MAC [38], CD-MAC [39] and many more. Most of them are feasible enough to be deployed in hardware. Since most of the protocols are contention based alone, so they use IEEE 802.11 DCF mode as the starting point for their research.

Most protocols covering this topic are not straight forward but they help in either reducing energy consumption or increasing diversity gain or selecting the best relay. Even then there are trade-offs to be considered in each research. This means that sometimes the relay nodes are beneficial for transmission but this consumes extra energy. So a decision has to be made between the energy required and relay transmission. Another considerable trade-off occurs when using multiple relay nodes for increasing diversity but it actually results in a higher collision probability.

In a rapidly changing environment, these above stated trade-offs are even more. In other words, the energy of overhead incurred and relay selection is measured according to the needs and possible output of the protocols. It nonetheless calls for further research. Efficient algorithms for choosing optimal helper from potential helper nodes also propose an open research field.

Moreover, most of the protocols use only a single relay either as backup for the failed direct link or as a helper for the source. Increment in their number can open a way for more diversity gain resulting in an increased overhead. Hence, keeping the overhead to a minimum and using more than one relay nodes for cooperation also pose a research question. Mobility and security are extremely relevant research problems in the cooperative scenario. All these issues have to be solved properly so that a high performance MAC with cooperation can be implemented in practice.

REFERENCES

- [1] Subir Kumar Sarkar, T. G. BASavaraju, Puttamadappa. "Ad-Hoc Mobile Wireless and applications: Principles, Protocols and Applications", Auerbach Publications.
- [2] John L. Hennessy, David A. Patterson, "Computer Architecture : A Quantitative Approach", MK Publications, fifth Edition.
- [3] Nosratinia, A, Hunter, T.E. ; Hedayat, A. "Cooperative communication in wireless networks", Oct. 2004, 0163-6804
- [4] Al-Karaki J.N, Zarqa, Jordan, Kamal, A.E. "Routing techniques in wireless sensor networks: a survey", Dec. 2004, 1536-1284

- [5] Hong, Yao-Win, Wan-Jen Huang, Fu-Hsuan Chiu, Kuo, C.-C.J. "Cooperative Communications in Resource-Constrained Wireless Networks", May 2007, 1053-5888.
- [6] Marco Contia, Song Chongb, Serge Fdidac, Weijia Jiad, Holger Karle, Ying-Dar Linf, Petri Mähöneng, Martin Maierh, Refik Molvai, Steve Uhligj, Moshe Zukermada, Via G. Moruzzi. "Research challenges towards the Future Internet; Computer Communications", Volume 34, Issue 18, Pages 2115–2134, 1 December 2011,
- [7] Scaglione, A., Yao-Win Hong, "Opportunistic large arrays: cooperative transmission in wireless multihop ad hoc networks to reach far distances", AUGUST 2003, 1053-587X.
- [8] Papadimitratos, Panos , Haas, Z. J., "Secure Routing for Mobile Ad Hoc Networks", 193-204, 2002.
- [9] Laneman, J.N., Wornell, Gregory W., " Cooperative diversity in wireless networks: Efficient protocols and outage behavior", 0018-9448, December, 2004
- [10] Murat Uysal, "Cooperative Communications for Improved Wireless Network Transmission; Information Science Reference", 1st Edition.
- [11] Sunmyeng Kim, "Relay Node Selection Algorithm for Cooperative Communications in Wireless Networks", International Journal of Computer, Control, Quantum and Information Engineering Volume 8, No:3, 2014
- [12] Lin Dai, Khaled B. Letaief, "Throughput Maximization of Ad-hoc Wireless Networks Using Adaptive Cooperative Diversity and Truncated ARQ", IEEE Transactions on Communications, Volume 56, no. 11, November 2008.
- [13] Kun Xie, Hui Li ; Xin Wang ; Shiming He ; Jigang Wen ; Guizani, M., "Routing and channel assignment in wireless cooperative networks", Wireless Communications and Mobile Computing Conference (IWCMC), Page(s): 1188 – 1193, 4-8 Aug. 2014
- [14] Aggelos Bletsas, Ashish Khisti, David P. Reed, Andrew Lippman, "A Simple Cooperative Diversity Method Based on Network Path Selection", IEEE Journal on Selected Areas in Communications, Volume 24, NO. 3, March 2006.
- [15] YANG Wei, LI Li-hua, SUN Wan-lu., "Energy-Efficient Relay Selection and Optimal Relay Location in Cooperative Cellular Networks with Asymmetric Traffic", The Journal of China Universities of Posts and Telecommunications.
- [16] Weihua Zhuang and Yong Zhou, "A Survey of Cooperative MAC Protocols for Mobile Communication Networks", Journal of Internet Technologies, Vol. 14 No. 4, P.541-560 (2013/7).
- [17] Lalitha Sankaranarayanan, Gerhard Kramer, Narayan B. Mandayam, "Cooperative Diversity in Wireless Networks: A Geometry-inclusive Analysis", IEEE Conference on Communication, Control and Computing, Pages 1598-1607, 2005/10
- [18] Thanasis Korakis, Zhifeng Tao, ShashiRaj Singh, Pei Liu, Shivendra Panwar, "Implementation of a Cooperative MAC Protocol: Performance and Challenges in a Real Environment", Journal on Wireless Communications and Networking 2009,
- [19] Unnikrishnan M, Rinu Titus, Premkumar C., "MAC Protocols for Cooperative Diversity in WLAN", International Journal of Research in Engineering and Technology Volume: 03 Special Issue: 15, Dec-2014
- [20] Marco Luise, Silvano Pupolin, "Broadband Wireless Communications: Transmission, Access and Services", Springer 1998, pg 255- 232.
- [21] Nansai Hu, Yu-Dong Yao, Zhuo Yang. "Analysis of Cooperative TDMA in Rayleigh Fading Channels", IEEE Transactions on vehicular technology, Volume:62, page(s):1158 - 1168 , March 2013.
- [22] Elzbieta Beres, Raviraj Adve., "Cooperation and Routing in Multi-Hop Networks", ICC 2007: 4767-4772.
- [23] Ying Zhu, Qi Han., "The More Relay Nodes, The More Energy Efficient?", International Conference on Computational Science and Engineering, 2009, Volume:2, Pg 87 – 92, Aug. 2009
- [24] Lei Jiang, Lars Thiele and Volker Jungnickel, "Modeling and Measurement of MIMO Relay Channels", IEEE VTC, page 419-423, 2008
- [25] Y. Zhou, J. Liu, L. Zheng, C. Zhai, and H. Chen, "Link-utility-based cooperative MAC protocol for wireless multi-hop networks", IEEE Transactions Wireless Communications, vol. 10, no. 3, pp. 995–1005, Mar. 2011.
- [26] C. Zhai, J. Liu, L. Zheng, and H. Xu, "Lifetime maximization via a new cooperative MAC protocol in wireless sensor networks", IEEE GLOBECOM, Nov. 2009.
- [27] N. Marchenko, E. Yanmaz, H. Adam, and C. Bettstetter, "Selecting a spatially efficient cooperative relay", IEEE GLOBECOM, Dec. 2009.
- [28] R.B. Groenevelt, E. Altman, P. Nain, "Relaying in MobileAd-Hoc Networks", WiOpt 2004.
- [29] Mohaisen, R, Al-Mistarihi, M.F. , Sharaqa, A., "Closed-form expression for outage probability in relay-based cooperative diversity systems over Nakagami fading channels with interference", International Convention on Information and Communication Technology, Electronics and Microelectronics, Page(s): 577 – 581, May 2014
- [30] Thanasis Korakis, Zhifeng Ta, Yevgeniy Slutskiy, Shivendra Panwar, "A Cooperative MAC protocol for Ad Hoc Wireless Networks", Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom, 2007
- [31] Murad Khalid, Yufeng Wang, Ismail Butun, Hyung-jin Kim, In-ho Ra and Ravi Sankar, "Coherence time-based cooperative MAC protocol for wireless ad hoc networks", EURASIP Journal on Wireless Communications and Networking, 2011:3, 2011
- [32] Hongzhi Jiao and Frank Y. Li., "A TDMA-based MAC protocol supporting cooperative communications in wireless mesh networks", International Journal of Computer Networks & Communications (IJCNC) Vol.3, No.5, Sep 2011.
- [33] Zhuo Yang, Yu-Dong Yao ; Xiaochen Li ; Di Zheng, "A TDMA-based MAC protocol with cooperative diversity", IEEE Communications Society, Volume:14 , 6, Page(s): 542 - 544
- [34] Ting Zhou, Hamid Sharif, Michael Hempel, Puttipong Mahasukhon, Wei Wang, Tao Ma, "A Novel Adaptive Distributed Cooperative Relaying MAC Protocol for Vehicular Networks", IEEE Journal on Selected Areas in Communications, Volume 29, no. 1, January 2011.
- [35] Jin Zhang, Qian Zhang, Weijia Jia. "VC-MAC: A Cooperative MAC Protocol in Vehicular Networks", IEEE Transactions on Vehicular Technology, Volume 58, no. 3, March 2009
- [36] Rasel Rana, Faruque Hossain, Kamrul Hasan, Uzzal Hossain, Obaidur Rahman, "Distance-Aware Reliable Relay Selection Provisioning in Wireless Cooperative MAC Protocol", 978-1-4799-6062-0/14/2014 IEEE

- [37] Devu Manikantan Shila, Tricha Anjali and Yu Cheng, "A Cooperative Multi-Channel MAC Protocol for Wireless Networks",
- [38] Aytac Azgin Yucel Altunbasak Ghassan AlRegib, "Cooperative MAC and Routing Protocols for Wireless Ad Hoc Networks", IEEE Communications Society.
- [39] Sangman Moh, Gwangju, Chansu Yu, Seung-min Park, Heung-Nam Kim, "CD-MAC: Cooperative Diversity MAC for Robust", International Conference on Communication in Wireless Ad Hoc Networks, 2007.

Architectural Framework for Inter-Agency Intelligence Information Sharing Knowledge Based System in Botswana

Ezekiel U Okike and T. Leburu-Dingalo

Computer Science Department, University of Botswana
Gaborone, 00000, Botswana

Abstract

The 21st century is witnessing increasing wave of terrorist and crime activities. This has made the need for intelligence-led policing the focus of most national governments including Botswana. The main objective of this paper is to propose an architectural model for intelligence-Led policing in Botswana which will provide support for all intelligent agencies in the country namely the Botswana Police Service (BPS), Directorate of Intelligence and Security (DISS), Directorate of Corruption and Economic crime (DCEC), and Criminal Investigation Department (CID). The model provides for inter agency information sharing using appropriate modern technologies in other to enable the agencies access to useful crime data which will enhance their ability to make informed decisions and take appropriate actions against cyber and economic crimes.

Keywords: *Cyber and economic crime, Security, intelligence policing, architectural framework, Knowledge base, Information sharing, system model.*

1. Introduction

One of the challenges facing the law enforcement is the detection and prevention of crime which has rapidly increased in recent years especially with the advent of cybercrime. Today's crime has far more reaching consequences and has a potential to affect more victims due in part to criminals access to sophisticated tools and large number of victims at a given time through the Internet. For instance an attacker who manages to gain access to a bank's database not only steals customers' information but also other organizational valuable

information and money. Thus it is crucial for law enforcement to prevent a crime from being committed as opposed to apprehending criminals after a crime has been committed.

Information is a crucial organizational resource comparable in value to land, labour and capital (Diwan, 2002; Okike and Mbero, 2015). Its importance in the detection, prevention and prosecution of crime has also been recognized in the law enforcement (Holgersson, Gottschalk and Dean, 2008). In recent years crime prevention methods have focused on approaches that are geared towards the effective and efficient utilization of information within the law enforcement field (Weiburd and Groff, 2015; Braga, 2015; Plant and Scott, 2009; Weiburd and Eck, 2004). One such approach stresses the gathering of intelligence information which can then be shared among various policing agencies. A significant example of a strategy that takes advantage of this approach is intelligence-led policing (, U. S Department of Justice (BJA), 2005, 2008; Texas Department of Civil Defence; ESRI; Hengst, M and Mors, J; wiki.)

Intelligence-led policing is defined as "an information organizing process that allows police agencies to better understand their crime problems and take a measure of the resources available to be able to decide on an enforcement or prevention strategy best designed to control crime" (Ratcliffe and Guidette [2008]) This model can thus be seen as advocating for proactive approaches that enable decision makers to know about potential crimes, which enable them to make well informed decisions and plans about detecting and preventing these crimes.

2. Problem statement

In recent times, crime in Botswana has increased significantly and often with more complex and sophisticated approaches from the criminals. The target of the criminals are individuals and corporate establishments with potential to adversely affect the national economy if urgent steps are not taken to mitigate cyber and economic crimes in the country. The Government of Botswana has also been in the process of implementing her national cyber crime policy aimed at addressing cyber crime issues in the country.

3. Aims and objectives

The main objective of this paper is to propose an architectural model for intelligence-Led policing in Botswana which will provide support for all intelligent agencies in the country namely the Botswana Police Service (BPS), Directorate of Intelligence and Security (DISS), Directorate of Corruption and Economic crime (DCEC), and Criminal Investigation Department (CID). The model provides for inter agency information sharing using appropriate modern technologies in order to enable the agencies access to useful crime data which will enhance their ability to make informed decisions and take appropriate actions against cyber and economic crimes.

4. Intelligence-Led Policing and Agency Information Sharing

Intelligent-led policing coupled with inter-agency information sharing is a good approach in tackling cyber and economic crime scenarios. According to Molokomme (2011), the ramifications of many cybercrimes are so extensive and technologically complex that they require specific knowledge to better understand the evolving nature of the threats as well as to develop successful preventative and/or investigative tactics and strategies. It is therefore crucial that the policing workforce should be properly trained and knowledgeable regarding this type of crime.

In Botswana there are four main policing agencies namely Botswana Police Service (BPS), Directorate of Intelligence and Security Services (DISS), Directorate on Corruption and Economic Crime (DCEC), and Criminal Investigation Department (CID). These agencies each have their special purpose as far as policing is concerned. Subsequently each agency will have their own type of information regarding cybercrime detection, prevention and investigation. However the idea behind addressing cybercrime is finding

proactive ways to identify potential threats and prevent attacks from taking place.

As policing agencies work with sensitive information a lot of challenges are expected as far as the storage of data, as well as the access and exchange of such data is concerned. However with modern information technologies, principles and tools these challenges can be minimized enabling effective and secure information sharing.

A technological system that will allow storage of intelligence data while at the same time providing operations that support effective information and knowledge sharing within the different policing agencies as being proposed in this paper must essentially be an expert system. Kendal & Green (2007) describe in detail the capabilities of such a system. The system has the ability to emulate the decision-making capabilities of a human expert. Thus an expert's knowledge can be computerized and made available to several users at the same time. This provides for more effective sharing of knowledge as opposed to an expert being consulted individually or agencies working as individual units. In addition, there are several knowledge representation methods used in the systems such as production rule, semantic net and frame. The system uses the production rule method as well as series of IF-THEN rules to deduct a choice based on some data.

The 21st century has witnessed increase in terrorist activities as well as in cyber and economic crimes. Due to this, there has also been a shift towards increased counter terrorism responsibilities for the police and other intelligent agencies. As expressed in [USDoJ], this requires the integration of both community policing and counterterrorism and depends strongly on community relationships. Hence intelligence led policing provides the much needed strategic integrating of intelligence with the following overall benefits [ibid]:

- Information and intelligence sharing among intelligence agencies
- Access to neighbourhood information
- Information management
- Two way communication with the public
- Scientific data analysis
- Problem solving
- Immediate threat management
- Encourage prevention of crime through proactive policing
- and more

The need for changes in law enforcement intelligence model became very necessary after the September 2001 terrorist attack in the US. The US government immediately

reviewed and studied her intelligence operations with a view to inculcating appropriate changes and needed improvements. Intelligent infrastructures were given needed recognitions at the federal level while similar efforts focused at enhancing local law enforcement intelligent operations such that the law enforcement organs play pivotal roles in homeland security (BJA, 2008). Similarly, national intelligent agencies of other nations have active responsibilities in ensuring security against terrorist threats and crimes in their national domains. A new approach to effective intelligence operations revolves around models that provide quick (real time) access to information among intelligent agencies and the public

5. System Models and Architecture

Figure 1 shows an expert system component upon which the system architecture is based.

5.1 Proposed System Architecture

An architecture framework deals with the design of a high-level structure for an information system. It results from assembling a certain number of architectural elements in some well-chosen forms to satisfy the major functionality and performance requirements of an information system, as well as some other, non-functional requirements such as reliability, scalability, portability, and availability (Kruchten, 1995)

The basic components for this system are the user interface, knowledge base, inference mechanism and the knowledge acquisition model as shown in Figure 2 and described below:

1. Explanatory Access Interface
 - a. Allow actors to interact with the knowledge management interface
 - b. Provide information retrieval and input
 - c. Provide reasoning for conclusion or response rendered
2. Knowledge Base
 - a. Stores information, data and facts
 - b. Rules to infer knowledge
3. Inference Mechanism
 - a. reasoning mechanism, that will be used to manipulate the stored knowledge and produce solution or suggestion to a problem
4. Knowledge Acquisition Module

- a. Used to assist in the development of the knowledge base

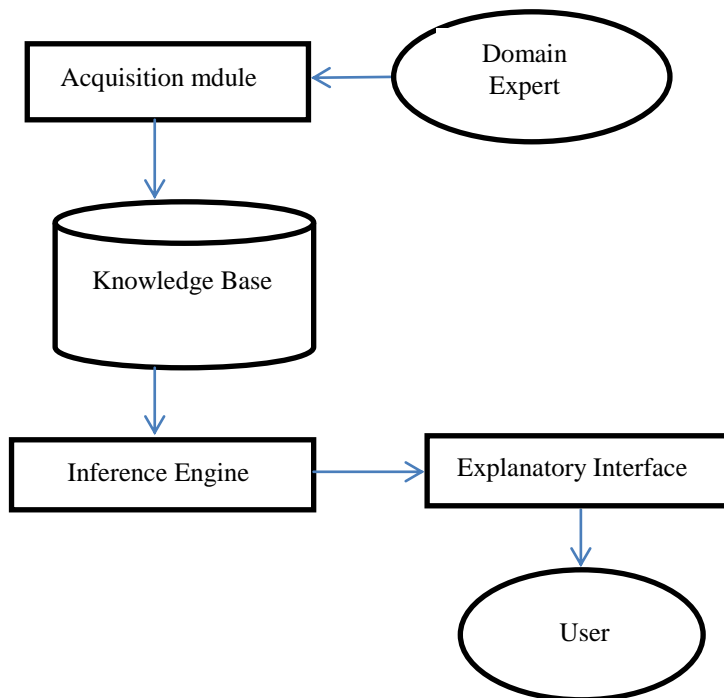


Figure 1: Expert System Components

5.2 Components of the proposed system

User Interface

System has two forms of interfaces, Expert and Normal User. These user interfaces are dependent on the target users of the actual system and their tasks as shown in the use case diagram in Figure2.

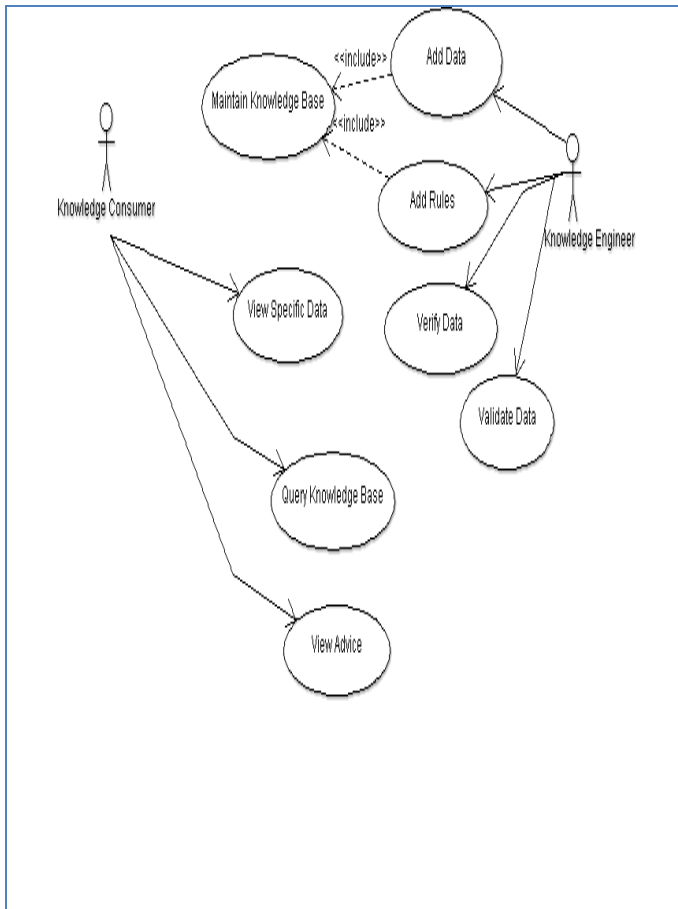


Figure 2: Use Case Diagram

Inference Engine

The engine uses a rule-based reasoning method with forward chaining. The methods entail extracting keywords from the users query, comparing these with the encoded data in the knowledge base then using a series of rules to generate a response.

Knowledge Base

Law enforcement officers are often times inundated with loads of information related to crime thus knowledge management is a crucial aspect of policing especially when it comes to cybercrime. This system is aimed at addressing this challenge. However since knowledge based systems are meant to address specific domains, the knowledge base for this system will be restricted to facts and information on cybercrime in the context of Botswana. In developing the knowledge base for the system the following information items can be used:

- Reported Incidences of Cyber crime
- Reported Cases and Resolution of Cyber crime
- Potential threats of cyber attacks

The domain facts and data should be derived from available documentation, interviews with law enforcement officers, citizens, stake holder companies such as banks and information technology companies.

Knowledge Acquisition Model

Since the data and facts for the knowledge base will be collected from various agencies, a knowledge acquisition module is included in the architecture to assist the knowledge engineers in standardizing and formalizing data and facts before encoding into the knowledge base. The module will also be used to validate and verify the data and check if it does not appear in the existing database.

The proposed system architecture is presented in figure 3 below.

1. Expert interface

- a. Authenticate and verify the identity of the user
- b. Route user requests to a module which can be used to create, update or modify the knowledge base or rule base
- c. Route the requests to a module which can be used to modify the programs used by the reasoning mechanism

2. Normal User

- a. Open a module that will verify the identity of the user
- b. On authentication allow access to the user to either
 - i. View details of cases, incidents, potential Threats or
 - ii. Type in a query requesting advice or solution to a problem

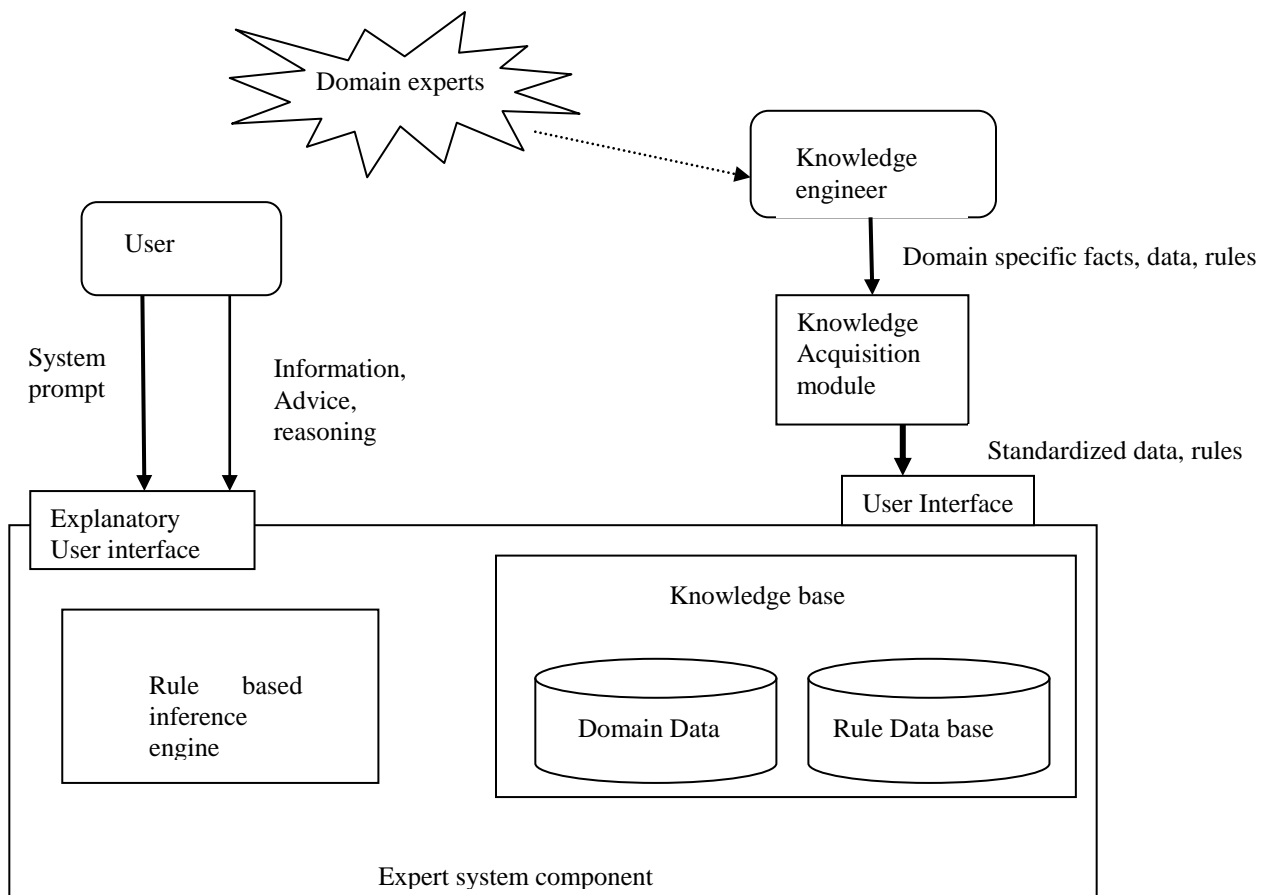


Figure 2: Proposed system Architecture

5.3 Inter Agency Information Sharing Data Flow Diagram

Figure 4 below shows the Data Flow Diagram (DFD) for inter agency intelligence information sharing

5.4 Intelligence-Led Policing Entity Relationship (E-R) Diagram

Figure 5 below provides a generic intelligence –led policing E-R diagram.. The four intelligence agencies are identified and represented. For each agency, all necessary entities must be identified with their relationships, and then finally related with other agencies. Figures 6-10 represents schemas for police service department only.

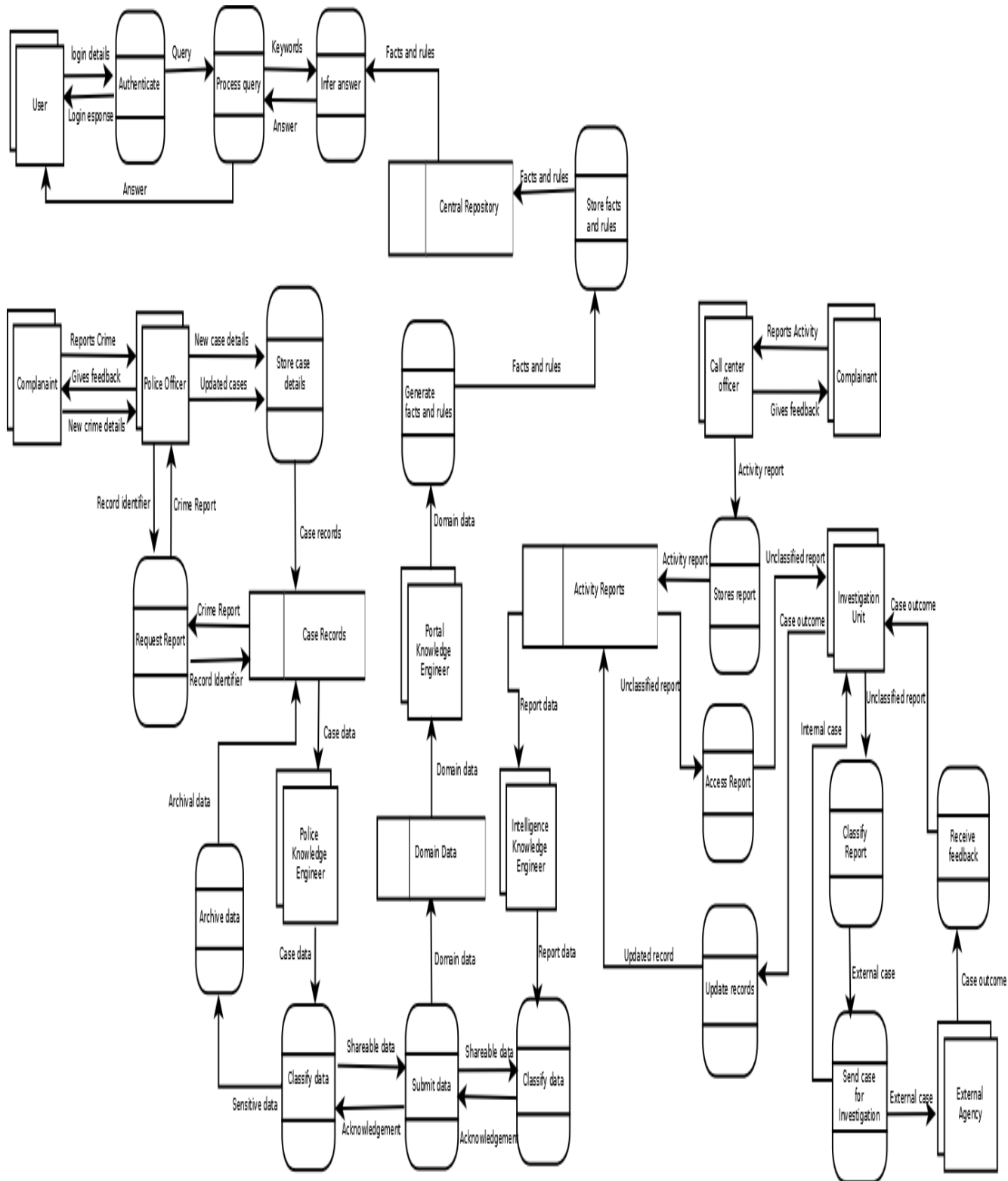


Figure 4: Data Flow Diagram for Inter agency information sharing

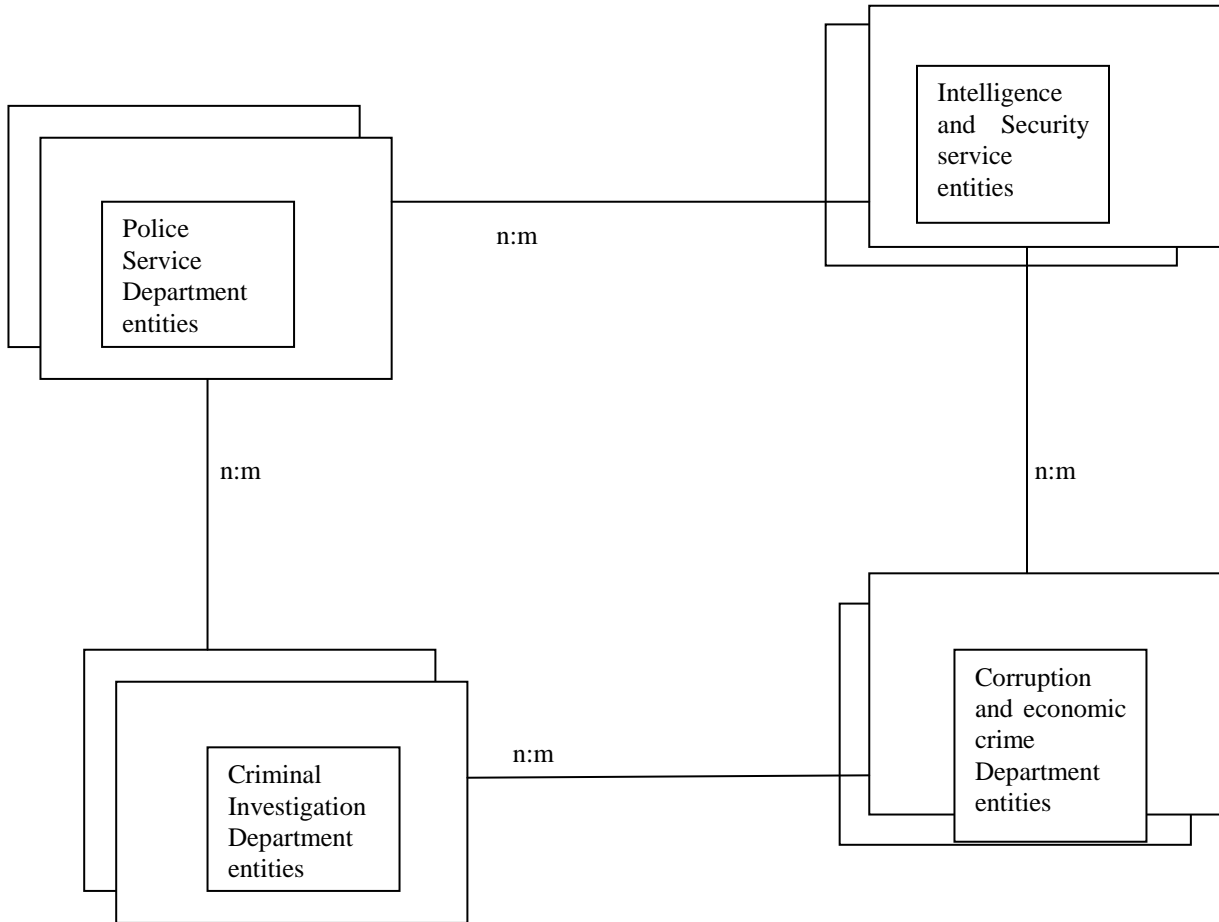


Figure 5: Generic Inter Agency E-R Diagram

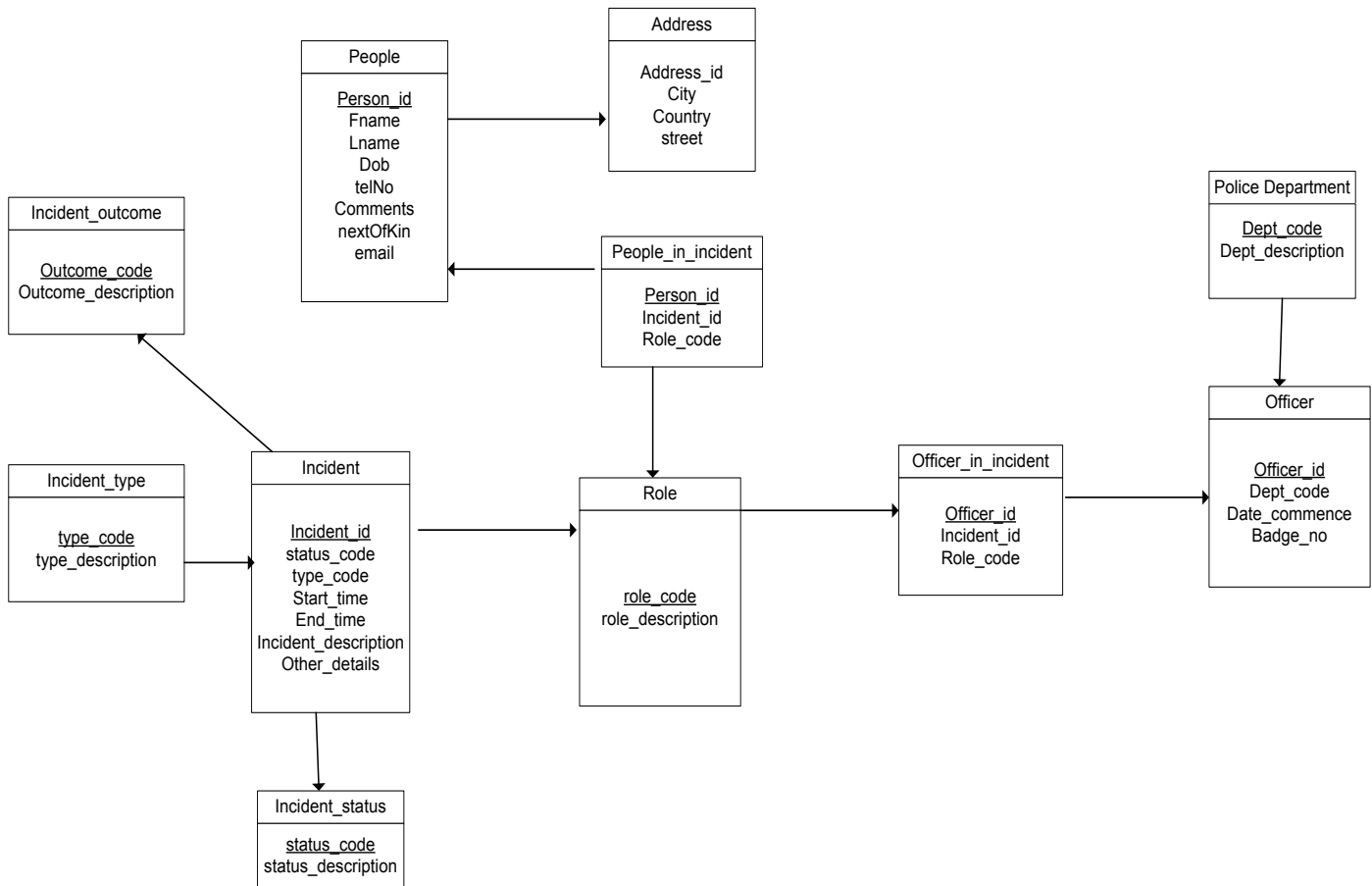


fig 6: Schema representing tables, attributes and relations of community policing

Step 1_Identification of Entities

- People
- Officers
- Incident
- Roles
- Police Department
- People_in_incident
- Officers_in_incident
- Address
- Incident_type
- Incident_status
- Incident_outcome Identification of Relationships

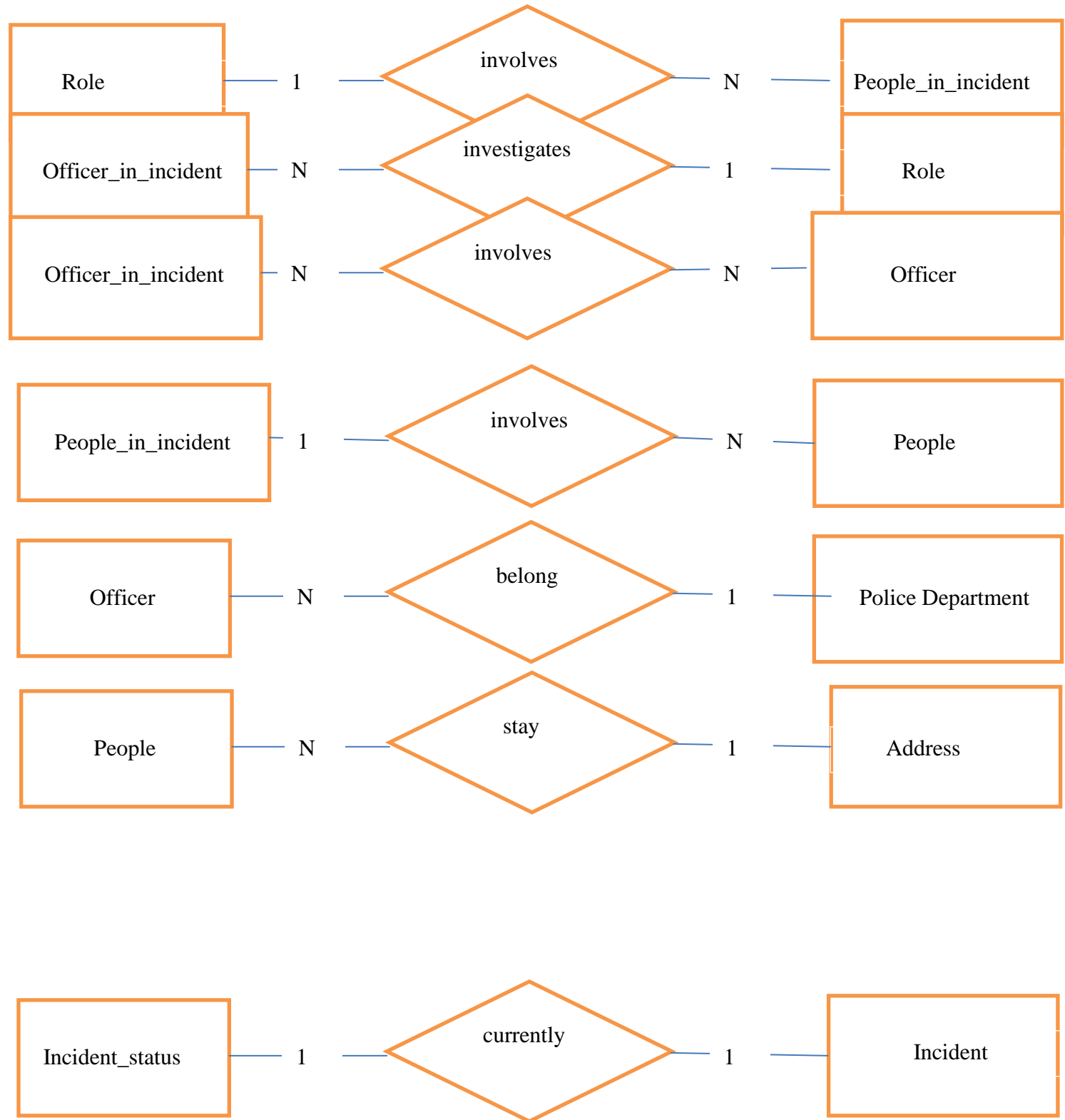
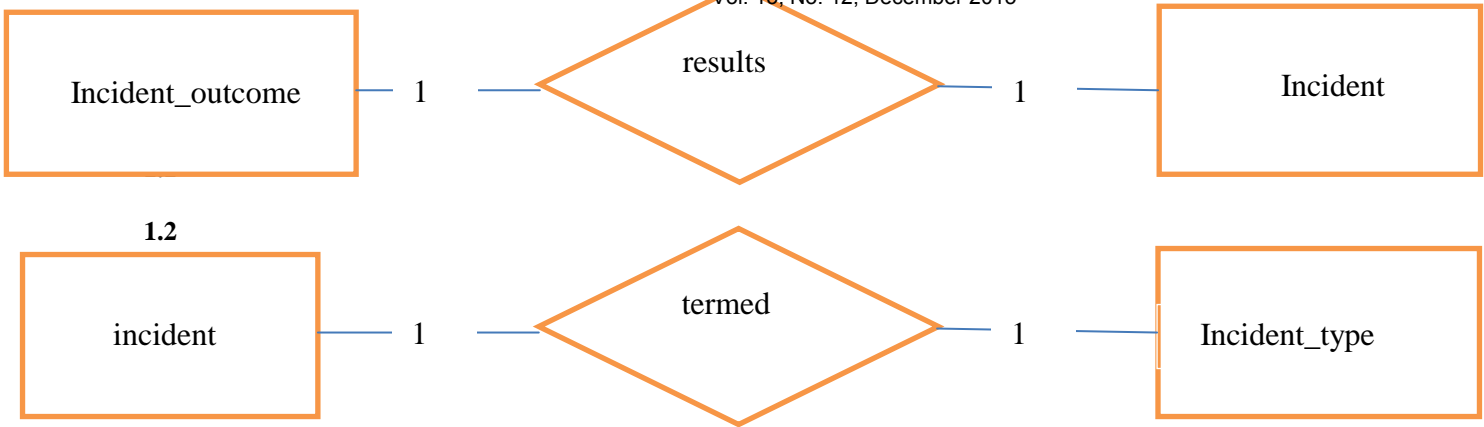


Figure 7: Community Policing E-R Diagram



Step 2_ Identifying key attributes

- People(**person_id**)
- Officer(**officer_id**)
- Incident(**incident_id**)
- Police Department(**dept_code**)
- Role(**role_code**)
- Police Department(dept_code)
- People_in_incident(incident_id)
- Officers_in_incident(officer_id)
- Address(address_id)
- Incident_type(type_code)
- Incident_status(status_code)
- Incident_outcome(outcome_code)

Step 3: Identifying other attributes

- People(fname,lname,dob,telNo,nextOfKin,comment,email)
- Officer(badgeNo,date_commence)
- Incident(incident_description,start_time,end_time,other_details)
- Police Department(dept_description)
- Role(role_description)
- Police Department(dept_description)
- Address(city,country,street)
- Incident_type(type_description)
- Incidt_status(status_description)
- Incident_outcome(outcome_description)

E-R Diagram for community policing

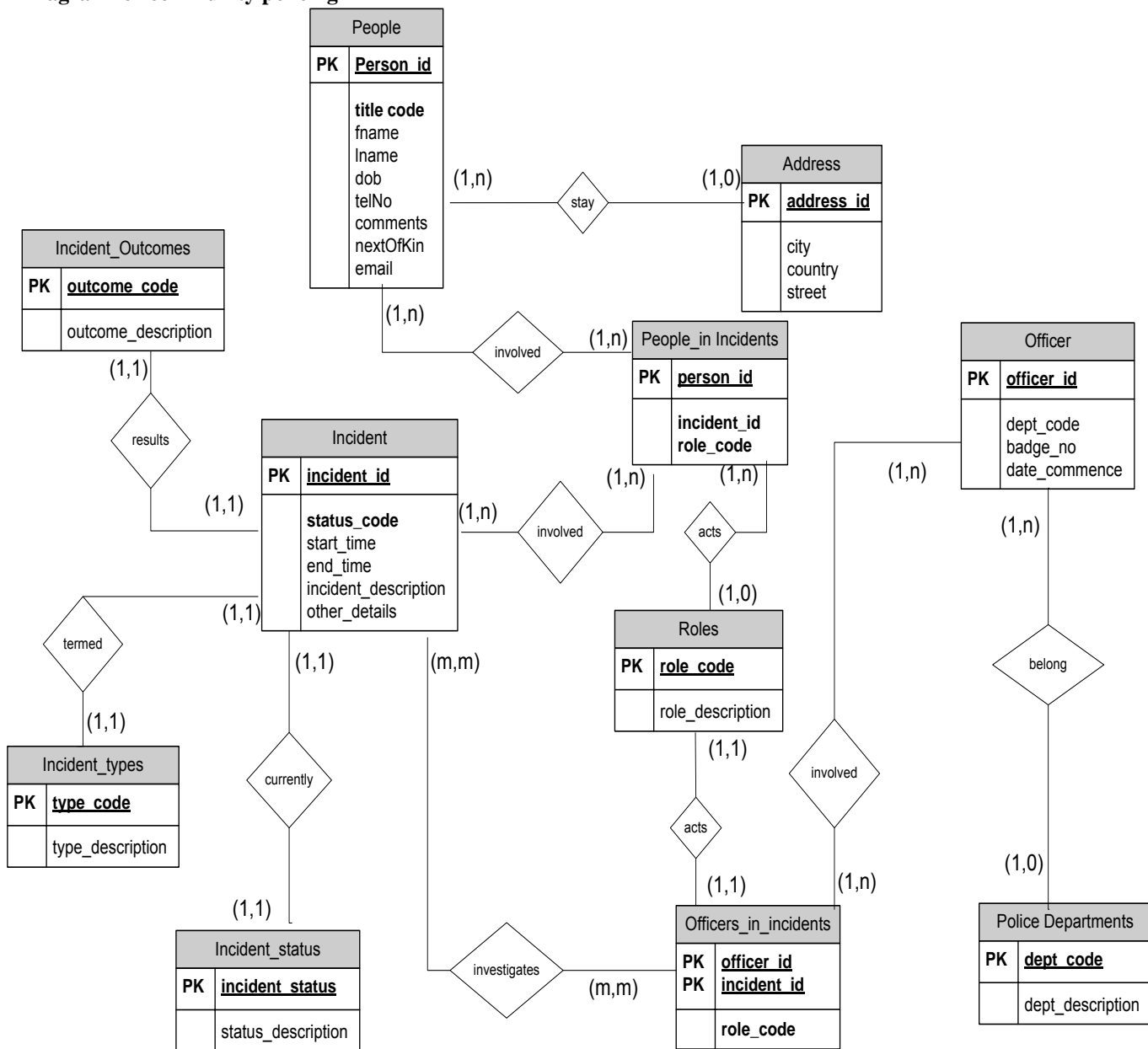


Fig 8. ER diagram for community policing

Transforming E-R diagrams into relations

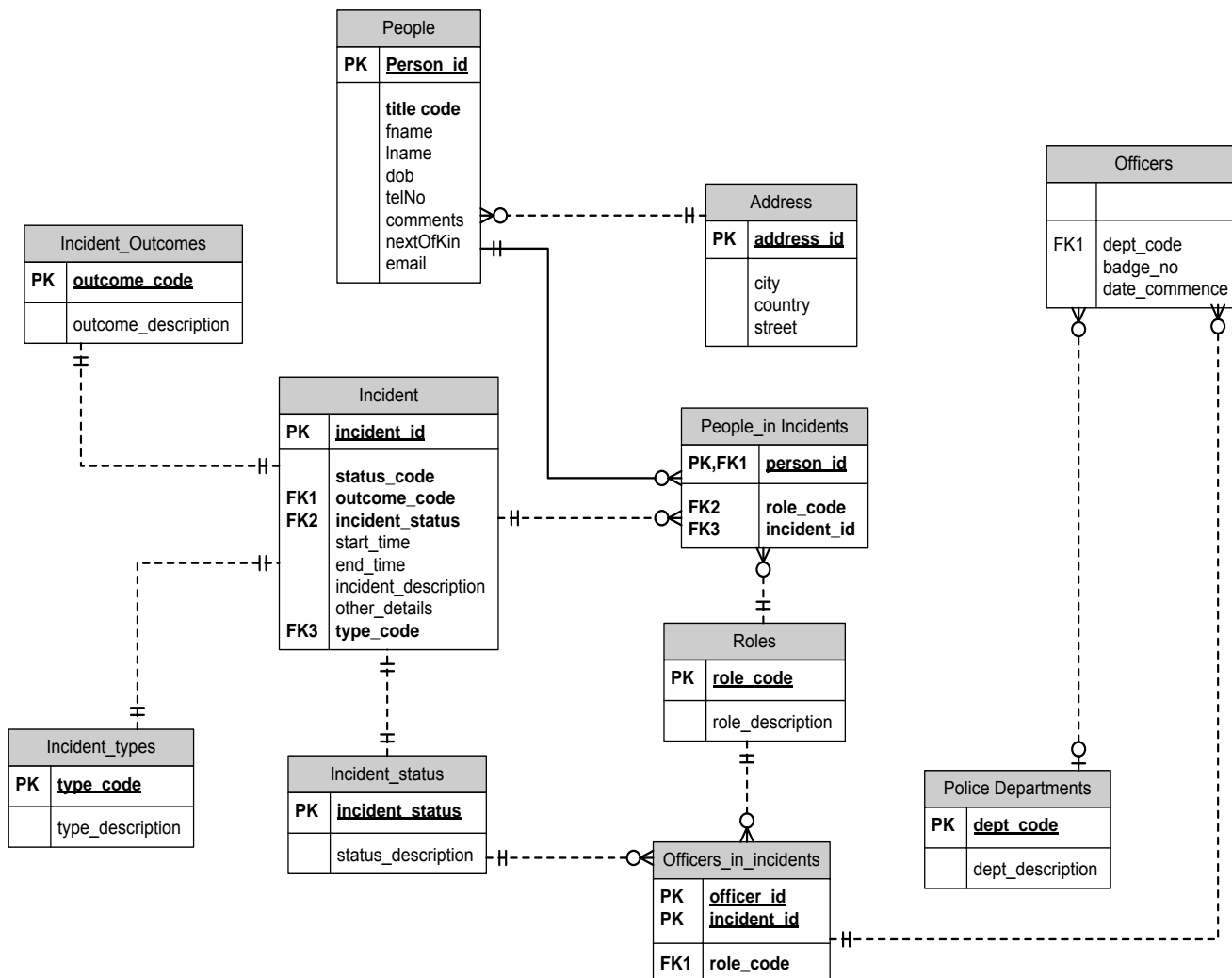


fig 9. Relation model of community policing

Relational schema notation in community policing

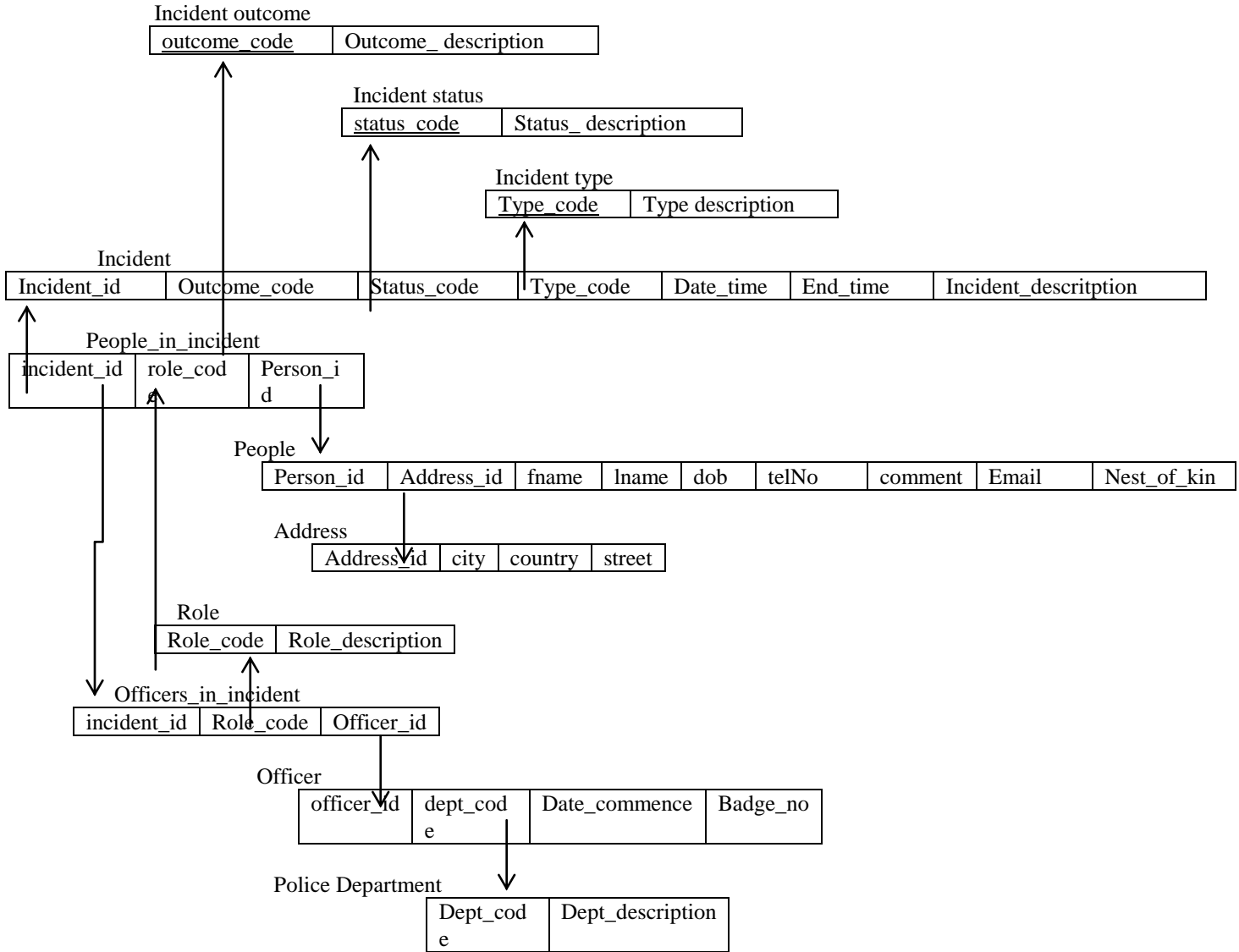


Fig 10. Schema notation illustrates how the tables are related even with distant tables.

6. Conclusion

The need for an intelligence-led policing framework for Botswana cannot be overemphasized. Our proposed architectural model may be further investigated for practical workability within the present set up of intelligence agencies in Botswana. The core of the system is the information base which is centralized for all the agencies. The model also emphasize quick response to intelligence information gathering and sharing by real time information alerts on the systems of integrated intelligence policing.

7. References

- [1] Kruchten P.(1995) Architectural Blueprints View Model of Software Architecture
- [2] BJA (2008). Intelligent led Policing: The new intelligence Architecture. Bureau of Justice Assistance (BJA), USA
- [3] Braga, A. A (2015). Crime and Policing revisited. New Perspectives in Policing, USA: National Institute of Justice
- [4] Diwan, P (2002). Information System Management. Kuala Lumpur: Golden Books
- [5] ESRI ().GIS provides the Geographic Advantage for Intelligence-Led Policing. (http://www.esri.com/library/fliers/pdfs/intelligence-led_policing.pdf) retrieved 10/10/2015
- [6] Jerry H; Ratcliffe, J. H., Guidetti, R. (2008). State Police Investigative Structure and the Adoption of Intelligence-led Policing, Policing: An International Journal of Police Strategies & Management, Vol. 31 Iss 1 pp. 109 -128
- [7] Kendal S.L., Creen,M., “Types of Knowledge-Based Systems”, An Introduction to Knowledge Engineering, Springer, London (2007)
- [8] Hengst, M and Mors, J (?). Community of Intelligence: The secret behind Intelligence-Led Policing. (http://www.policeacadie.nl/.../lect0intelligence/.../PID_2397493.pdf) retrieved 10/10/2015
- [9] Holgersson, S; Gottschalk, P and Dean, G (2008). Knowledge management in law enforcement: Knowledge views for patrolling police officers. International Journal of Police Science and management, vol. 10 No.1, pp76-88
- [10] Molokomme, A., “Combating Cybercrime in Botswana”, University of Botswana Law Journal 13 U. Botswana L.J. (2011)
- [11] Okike, E. U and Mbero, Z. A. (2015). Information Technology (IT) Knowledge Management system stage model: A proposal for today’s workplace. International Journal of Computer Science Issues, volume 12, Issue 5, pp193-202
- [12] Texas Police Department Civil defence Battalion (?). Intelligence-Led Policing: The Integration of community Policing and Law Enforcement Intelligence (http://www.cops.usdoj.gov/pdf/e09042536_chapter_04.pdf) retrieved 11/11/2015
- [13] U.S Department of Justice Bureau of Justice Assistance (2005). Intelligence –Led Policing: The New Intelligence Architecture – New Realities Law Enforcement in the post 9/11 Era (NCJ 210681).
- [14] U.S Department of Justice Bureau of Justice Assistance (2008). Reducing Crime through Intelligence –Led Policing (2008-DD-BX-K675).
- [15] U.S Department of Justice Document (USDoJ). Intelligent led Policing: The Integration of Community Policing and Law Enforcement Intelligence. http://www.cops.usdoj.gov/pdf/e09042536-chapter_04.pdf
- [16] Weisburd, D and Groff, E (2015). The Dallas AVL Experiment: Evaluating the use of automated vehicle locator in Technologies in policing
- [17] Weisburd, D and Eck E. J (2004). What can police do to reduce crime, disorder and fear. Annals
- [18] Wiki (2015). Intelligence-Led Policing. (http://en.wikipedia.org/wiki/intelligence-Led_Policing) retrieved 10/10/2015



Dr Ezekiel Uzor Okike, MIEEE, MACM, MAIS is a Senior Lecturer in the Department of Computer Science, University of Botswana, Gaborone. He holds a PhD in Computer Science (2007), a Master of Information Science (M.Inf,Sc), 1994 and a B.Sc Computer Science, (1992) all from the University of Ibadan, Nigeria. He is a member of IEEE, ACM, and AIS. Presently he is the cluster chair of the Information Systems Cluster, Department of Computer Science, University of Botswana. He has published many papers in international journals and attended many international conferences where he presented papers. His research interests are in Software Engineering, Information Systems, Cyber Security, Programming languages, Compilers, and High Performance computing



Mrs T. Leburu-Dingalo is a Lecturer in the Department of Computer Science, University of Botswana. She holds a Bachelor of Engineering (BEng) from Florida Institute of Technology, and Master of Science (M.Sc.) degree in Information Systems from the University of Botswana. Her research and teaching areas are in Operating systems and Information systems.

A Comprehensive Review of Different Types of Cooperative MAC layer Protocols

Ferzia Firdousi

Department of Electrical Engineering
National University of Sciences and Technology
Pakistan

Abstract-For physical layer, cooperative systems have been devised which aim at increasing diversity gain achieved by spatial diversity. This advantage can be mapped onto the MAC layer to achieve throughput increment, faster transmission rate, reduction in power and energy consumption, and a large network coverage area. However in the race to achieve a good MAC layer protocol, many new problems are created, including a redundant use of relay nodes, increase in energy wastage and increase in delay etc. In order to understand the true capabilities of cooperative communication at the MAC layer, many protocols need to be studied and their drawbacks identified, so that the upcoming research can be improved. This paper clinches research regarding these different types of cooperative MAC protocols by summarizing and analyses them. The analytical result, discourses issues like which relay node is the optimal one, how can a system be made energy efficient and which methodology to be followed for forwarding a data packet in case of transmission failure, and explains them in detail to allow room for future, new and improved research.

Keywords: cooperative, MAC, relay selection, diversity gain, energy efficiency

I. INTRODUCTION

Wireless networks need to be energy efficient and have reduced delay in transmission time [1]. Medium Access Control (MAC) is one of the key areas where innovation in protocol can result in an energy efficient network [2]. It has the ability to eliminate the causes of energy wastage in the wireless network for example collision, overhearing, idle listening and overhead energy consumption due to control packets [3]. Overlap of two or more data packets during transmission results in collision [4], while overhearing occurs when a data packet not intended for a node is received by it [5]. Protocol overhead is incurred when nodes use control frames for channel reservation before data transmission; this consumes energy [6]. In idle listening, a node keeps its receiver on in the hope of receiving something even if the

channel has nothing for it; this also wastes energy [7]. MAC protocols can be modified to improve energy usage [8]. The only mandatory requirement while formulating a MAC protocol is that they must reduce or eliminate the energy wastage caused by some or all of the reasons stated above.

Provisionally, literature review of MAC protocols show that a large number of protocols proposed by different scholars work towards energy efficiency. They can largely be categorized into two groups; reservation based MAC protocols like TDMA and contention based MAC protocols like CSMA/CA in IEEE802.11, as explained below.

Not unlike a simple queued line depicting people who wait for their turn, nodes in reservation-based protocols follow a definite pattern of slot assignment in a channel on the basis of time, frequency or both. Accordingly the nodes transmit only within their assigned slot(s). Since everyone gets a turn in the channel for transmission therefore there is no collision. But the problem of clock synchronization for all the nodes comes into question. In addition reservation-based protocols cannot easily change with change in topology therefore they are highly incompatible and rigid in nature.

Unlike reservation-based MAC, contention-based protocols display nodes competing for access to the communication medium (channel) if they have a data packet they wish to transmit. This makes their working quite simple. This also means that none of the nodes need to be synchronized or have topologically correct information of the entire network for transmission [9].

With multiple data rates like 1, 2, 5.5 and 11Mbps in IEEE 802.11b wireless LANs [10], the diversity of the wireless channels is clearly visible. 1Mbps in the

system is the very basic data rate and provides the system with its maximum transmission range. Its advantage can also be seen with a resulting lower bit error rate (BER) as compared to the higher BER achieved with the other higher data rates therefore when the channel's quality is very poor, lower data rates are used to ensure a better transmission. But while it ensures a better transmission, it outputs a degraded network. Therefore in order to be able to use higher data rates, cooperative diversity is introduced [11].

The higher data rates are provided by the neighboring nodes of the source node i.e. the relay nodes, so that the data transmission from the source node with a lower data rate can be assisted by the neighboring nodes which can provide higher data rates for data transmission [12]. This is the main essence of cooperative diversity. It should be noted here that relay nodes should be free and compatible for cooperation [13]. It is highly desirable that these relay nodes provide data transmission at a higher data rate than the source node. But this does not always follow through so simply. This brings in the question of optimal relay selection. Thus a relay located beneficially can help improve the overall performance of the system [14].

In order to be able to apply cooperative diversity at the MAC layer, the layer and its protocol needs to be studied and researched thoroughly, so that the implementation becomes feasible. There is a lot of literature available on cooperative MAC layer protocols for this purpose. While most of these protocols are good, there is shortage of work which compares any of them in great detail. This is then the aim of the paper; it presents complete critical review of the available protocols compiled in a single place.

The rest of the paper follows this layout: section 2 gives a detailed review of the MAC protocols and their steps as proposed in the respective papers, section 3 works into a very comprehensive comparison of all the described protocols on the basis of the results and introduces further research ideas which still remain unanswered, and finally section 4 concludes the work.

II. COOPERATIVE MAC PROTOCOLS SCHEMES UP-TO-DATE

It is important that IEEE802.11 protocol be described here for comprehension purposes. In it [15], whenever a source node wishes to transmit a data packet using the channel, it senses the channel for idleness for duration of distributed interframe space (DIFS). If the channel is found idle, nodes back

off for duration of time usually called a contention window. After this window, the source node sends a ready-to-send (RTS) frame to a destination asking for permission to transmit. Destination, after duration of short interframe space (SIFS), replies back with a Clear-to-Send (CTS), but only if it is idle. This is the handshaking protocol. In order to avoid collision from other nodes while source is transmitting, the other nodes update their NAV table. This way they do not transmit for the duration of the ongoing transmission. After the transmission, source node awaits acknowledgment (ACK) response from the destination for a period of SIFS again. This completes one basic transmission cycle. The total duration of this handshake is $RTS+SIFS+CTS+SIFS+DATA+SIFS+ACK$.

The most common methods for cooperation are either contention based or reservation based MAC protocols. The most significant protocols and their working are described here. In each protocol, it is mentioned whether relay nodes are used as a backup link or as a link besides the direct link. Therefore during description it is clearly stated whether each protocol has a diversity gain of 1 or more. Ready-to-Send is labelled as RTS, Clear-to-Send as CTS and Acknowledgement as ACK.

A. CoopMAC

In this protocol described by [16], relay node is used when required by the network. The source decides whether to use indirect link or direct link for transmission. Therefore the diversity gain remains 1. It is a contention based system with modification in the RTC/CTS packets and addition of a few packets like Helper-to-Send (HTS). The metrics used for calculations of optimal r is the availability and reception of HTS packets from r on first come first serve basis. The process is summarized in Fig. 1.

CoopMAC modifies the IEEE802.11 DCF protocol and makes it more flexible since it allows source a choice between the direct and indirect data transmission links on the basis of the data rates of either links. Source nodes are free to choose their relays and can easily employ the relays' services for cooperation. All information is kept in a small table called the CoopTable as shown in Fig. 2; the data regarding different relay nodes is stored in the form of data rate of their links. The source only has to look into the CoopTable to make a decision regarding the links.

In order to maintain the CoopTable, a node overhears the ongoing transmissions at the other

nodes. It should be kept in mind that the data rate of any link between two nodes is calculated by Physical Layer Convergence Procedure (PLCP) header [17]. Therefore it can be assumed that S knows the data rate between itself and D from previous transmissions. Source also overhears and analyzes the transmission between r and D to decide whether that link is better than its own or not.

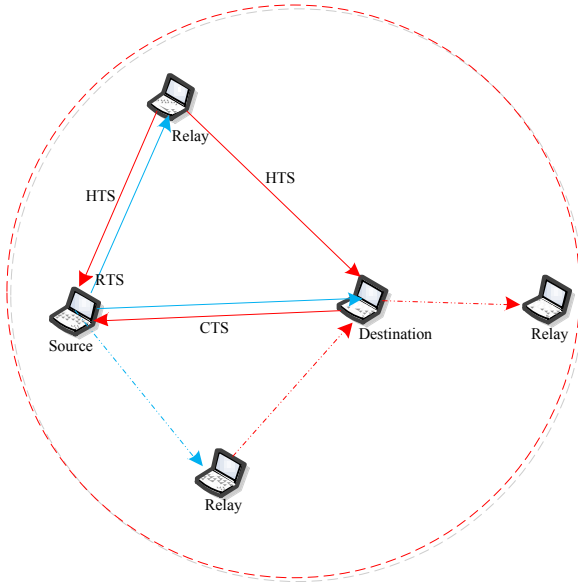


Figure-1: Protocol layout for finding available helpers in CoopMAC

Table 1: CoopTable for CoopMAC

| ID (48 bits) | Time (8 bits) | R_{Helper} (8 bits) | R_{Source} (8 bits) | $R_{\text{Destination}}$ (8 bits) |
|---------------------------|--|--|---|---|
| MAC addresses of helper 1 | Time the last packet heard from helper 1 | Transmission rate between helper 1 and destination | Transmission rate between source and helper 1 | Count of sequential transmission failures |
| | | | | |
| MAC addresses of helper N | Time the last packet heard from helper N | Transmission rate between helper N and destination | Transmission rate between source and helper N | Count of sequential transmission failures |

Conclusively, since for cooperation, cooperative handshaking is also needed, therefore an additional frame HTS is introduced. In case source decides to use the indirect link, it puts the address of both r and D in the RTS frame along with the request to

cooperate. Relay, if available, acknowledges this with an HTS. Finally, D responds with a CTS followed by the transfer of data.

B. CD-MAC:

This paper [18] renames the cooperative diversity in MAC protocol and calls it cooperative diversity (CD-MAC) protocol. It is a contention based protocol; control packets are used for conserving the channel beforehand. If channel contention fails, then an indirect link is employed along with direct link yielding a diversity gain of 2. In the other scenario, when source uses only direct link, then gain remains 1. System uses space time coding at the physical layer.

Since space time coding helps to send the signal via direct and indirect link together, whenever required, therefore, the main problem that remains is the selection of the r node. Optimal relay is chosen with the reception of a hello packet. This is accomplished by each node monitoring its neighbors and determining a single r as the one that exhibits the best link quality. Before making this choice three things are strictly contemplated by the source node: (i) is the communication between a node and its relay reliable. (ii) is the best relay most probably the closest node so the spatial diversity does not negatively impact cooperation (iii) do source and relay share the same communication environment? If all of these points are found for any relay node, it is used if cooperation is needed.

Source transmits an RTS to destination. If RTS successfully arrives at destination and destination is willing to receive then a CTS is transmitted which is received by source. But if CTS fails to be transmitted back then source retransmits the C-RTS to it destination along the direct and indirect link with the help of relay node. Destination again sends back the C-CTS along both the links. Reception of the C-CTS is the queue for the transmission of data packet which is also done via direct and indirect link. ACK again follows the same routine. Note that at relay, the received packet from source is first decoded before being re-encoded and forwarded to destination. This re-encoding is done in the second time slot 2 using DSTC.

The control packets including Cooperative-RTS (C-RTS) and similar ones are unmodified structurally. In other words they have the same structure as conventional RTS, CTS, DATA and ACK. The only difference seen is in the packets transmitted from source and relay. They are

structurally different at the MAC and physical layer due to space-time block codes.

It is seen that CD-MAC is performance-wise a lot better than the conventional DCF of the original IEEE 802.11, although this comes at the cost of increased end-to-end packet delay.

C. C-MAC

In [19] C-MAC stands for Cooperative MAC. Relay is preferred over direct link, but only if it replies back to source on time via a Ready to Relay (RTR) packet which is much like CTS packet. In addition, the control packets are modified drastically, forming CoopRTS/RTR/CTS. It is clearly a contention based system and the metric used for the selection of optimal relay circles around the relay which responds to the RTR on time.

This protocol is mainly based on IEEE802.11 protocol and is comparable to it. If no relay is used then its operation follows that of IEEE802.11 protocol but if an optimal relay is found then transmission changes drastically. Hence when possible it has a diversity gain of 2. The difference is created by the handshaking procedure which involves a four-way protocol instead of the usual three-way one for IEEE802.11.

CoopRTS is broadcasted by source to relay nodes offering them a cooperative link. RTR frame is used by relays to announce their willingness for cooperation. Based on the prior transmission knowledge, source uses either direct link or both direct and indirect link for data transmission. CoopRTS/RTR/CTS handshake procedure is used to ensure the participation of relay in the transmission procedure. If relay is willing to participate, source broadcasts the packet intended for destination so relays can also overhear this transmission. Destination keeps the signal received from source while waiting for the signal from relay node. Relay transmits the overheard signal from source to destination after a SIFS time. Combined decoding of these two signals is performed at the physical layer. Destination responds with a positive ACK feedback.

C-MAC protocol is imperfect because it lacks a good method of finding a potential relay, but it maintains backward compatibility with IEEE 802.11. It requires physical layer support for cooperation to jointly decode the received signals at destination. But, it provides diversity since both the direct and indirect link (if available) is used for transmission.

D. C-TDMA:

In [20] C-TDMA the all the nodes other than source and destination act as relay nodes. In this paper, each terminal has three tasks. Firstly, a user/terminal needs to transmit its own data packet during its allocated time slot in each frame. Secondly, it monitors/overhears other time slots in each frame. Finally, it cooperates with other terminals to retransmit their failed packets caused by channel fading/noise during previous frames. Therefore, a diversity gain can be achieved through these cooperative transmission procedures to increase the probability of correct packet reception. This paper models the C-TDMA protocol using Markov chains and derives the expressions of network throughput; packet dropping rate and average packet delay. This cooperative MAC design effectively overcomes severe channel impairments because of fading, increases the probability of correct packet reception, and thus results in throughput improvements. Although this paper formulates cooperative transmissions in case of TDMA MAC protocols which are known to conserve energy, the use of more than one relay nodes forms a very redundant system and thus wastes energy.

E. IrcMAC:

Instantaneous relay based cooperative MAC, IrcMAC for short, is the protocol that chooses the optimal relay as a backup transmission path [21]. The paper uses SNR estimates to calculate data transmission rates. Simplistically speaking, relay which offers a path that takes less time than the coherence time of the channel is the optimal relay. Since there is only one path used for transmission, therefore the diversity gain is 1. It is a contention based protocol, and RTS/CTS/ACK packets are the same as IEEE802.11. However, a new frame, called relay response frame (RRF), shown in Fig. is introduced in the paper which helps in estimating the best relay.

For the protocol, source senses channel for idleness, and if found idle, source sends an RTS packet to the destination. If CTS is received successfully from destination, source calculates the SNR for this link. Meanwhile, relay nodes use the reception of both RTS and the CTS packets to estimate the SNRs for the links of source-to-relay and relay-to-destination respectively and responds by sending an RRF back to source. This has information regarding a) SNR of source to relay link and b) SNR of relay to destination link. Responses from relays closer to source are received earlier than those from

other. This way, whichever response comes first, and has good SNR values, is chosen as the optimal relay

link for cooperation.

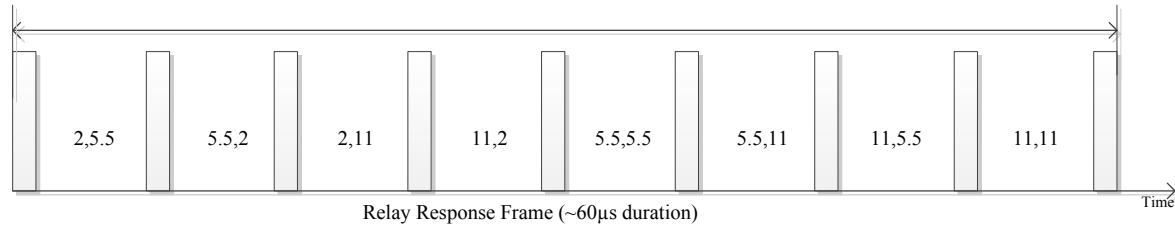


Figure 2: Relay Response Frame for IrcMAC

Once transmission via relay link is completed and destination receives the packet successfully, ACK is transmitted to source. If no ACK is heard from destination, node even after back up cooperation then source is left with nothing but to repeat the transmission cycle by resending the packet itself after a small waiting period.

F. VC-MAC

Vehicular cooperative or VC-MAC is the name coined for nodes which communicate via an access point (AP) due to their mobile nature [22]. Mobile nodes are always in contact with AP, but in cases when any of the nodes cannot reach AP, cooperative communication is used to forward access to AP for them by other nodes. This concept is illustrated in Fig. this paper raises some very interesting points; a) it is seen that here the sending/receiving nodes are the ones acting as relay nodes as well, b) cooperation is used not just for forwarding data, but is also used for access to Aps. The diversity gain of the protocol remains 1.

VC-MAC explores the spatial reusability in broadcast situations. There is also a necessary access point and all the vehicles are in contact with this point for their information. Although there are plenty of potential relay nodes only one is used to avoid collision. If a lot of relays were used then the spatial interference would increase and reusability will not be possible.

The system follows IEEE802.11 protocol. Four steps followed are namely, gateway broadcast period, information exchange period, relay set selection period and data forwarding period. As the name suggests broadcasting of the packets belonging to the gateway routers is done during the gateway broadcast period. This broadcast takes place in the upper layers going out to all vehicles in the area; information exchange is used by the potential relays to let other

nodes know about their existence; in the relay set selection period the best relay is selected from all potential relay nodes in a distributed manner by leveraging the back off mechanism; and selected relay broadcast the packets received from the gateway to potential destinations within its transmission range in the data forwarding period.

G. CoopMC MAC

This paper [23] modifies cooperative MAC protocol to Cooperative Multi-channel MAC (CoopMC MAC) protocol. As the name suggests the protocol is designed to enable the nodes including source, destination and the corresponding relay(s) to use a number of channels to simultaneously communicate via cooperative diversity. The diversity gain becomes 2 or more than 2. RTS/CTS/ACK is replaced by CTU/CTU-R/CTU-ACK.

This paper aims at minimising the interference in cooperative networks channels. Therefore MAC design improves channel capacity. There are two stages at work in this protocol; control stage and data stage as shown in Fig. 3.

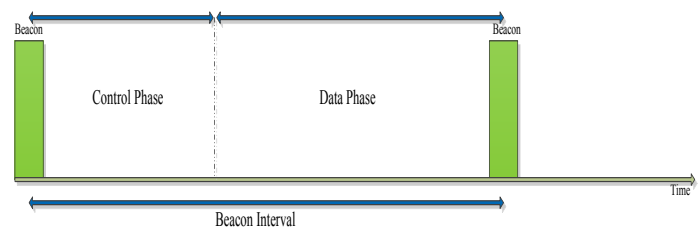


Figure-3: Protocol layout of CoopMC MAC

In control stage all relays that are willing to cooperate negotiate their position. Then the selected channels send the data to the destination in the data stage. In order to reduce the overhead, multiple data

packets are sent during transmission. The plus point of this protocol is that the transmission from source to destination, and consequently from relay(s) to destination, is done in a single time slot instead of two. The paper is not focused on selection of the best relay node alone. Instead it tries to find the best channel available as well.

Every node has a neighbour-table (NTable) which saves the data for the each of their one-hop neighbours. This helps in selection of the best relay. A table for the channel data (CTable) is saved by each node which has the data of the available channels for transmission. The channel for data transmission is selected on the basis of being used by the least available number of nodes in the vicinity of source. The rest of the protocol follows once the best channel has been selected.

According to the protocol, first a beacon is broadcasted for channel synchronization at the start of every time slot. Each node starts and ends at the same time. In the control phase each relay node which can send data to destination listens to the channel for collision. At the start of the protocol, source node waits for a DIFS period and then sends a channel-to-use (CTU) packet to relay nodes and destination. This CTU includes the MAC address of the optimal relay node and the corresponding channel information according to the CTable. Relay responds with a channel-to-use-Relay (CTU-R) packet which is also sent to destination. When destination first receives CTU and then a CTU-R, then it knows the channels through which data is going to be sent, therefore it sends a channel-to-use-Acknowledgment (CTU-ACK) back.

In this manner, two data packets are sent to destination. Specifically source sends a packet to relay and destination on one channel while relay uses a second channel to send data to destination. Destination acknowledges the transmission by sending back a CACK to source and relay.

The use of multiple channels is a very smart manner of increasing diversity gain while saving energy used. But at the same time, the paper seems to be convinced that the calculation of best channel can serve as a best metric for finding an optimal relay node. More precisely, the fact that a good channel found by the source, may eventually not be very good with respect to destination, is completely ignored. Bringing this factor into consideration can greatly influence the selection of best channel.

H. (DRRS) Cooperative MAC protocol

In [24] relay is selected on the basis of distance from both source and destination and SNR is also calculated by it. The diversity gain remains 1 and RTS/CTS are modified to HTS/HCTS in case of relay transmission. Distance and SNR represent the channel state information. If relay lies close to source but far away from destination it can give a false sense of optimality. Therefore for the distance and SNR metrics to work best, it should be kept in mind during the selection of relay that the criteria of lying just close to destination, alone, should not be fulfilled. Instead it should also have great SNR and lie between both source and destination.

In the protocol, source waits for DIFS and then sends RTS packets to destination. After receiving CTS, source sends the data packet to destination. This mode is called Direct Transmission Mode (DTM). However if CTS is not received then the Cooperative Transmission Mode (CTM) is initiated. In this mode source waits for a certain time period then sends a Helper to send (HTS) packet to potential relay nodes. If the relay nodes lie within the distance threshold (as set by source) they reply with a Helper-Clear-To-Send (HCTS) packet. This HCTS contains information regarding energy level and distance from the source. Thereafter, via relay transmission, the data packet is sent to destination and ACK is received.

III. ANALYSIS

Backward compatibility with IEEE802.11 has been maintained by all of these protocols therefore they are very easily implementable with current IEEE techniques. But there is still a lot of analysis regarding cooperation and their relays, left to be done.

Number of relay nodes, that are actually used for cooperation vary from 1 to many; whereas CoopMAC, CoopMC MAC, VC-MAC, DRRS-MAC, and IrcMAC use only 1 relay node for transmission, C-TDMA uses nearly all the other nodes as relay nodes.

In most cases relays are assumed to lie in between source and destination. In other words they are assumed to lie in the overlapping area of source and destination, which is not a very correct assumption, and only DRRS-MAC cares to mention it.

The question of whether relay nodes are used as back-up transmission path (cooperative relaying) or in cooperation with direct path (cooperative

diversity) comes next. Protocols find the most optimal relay and only engage it as a) a backup path if transmission via direct path fails like in CoopMAC and DRRS-MAC, or b) as path which acts as the link for transmission because transmission via direct path is not feasible like IrcMAC protocol. Only CD-MAC, DRRS-MAC and CoopMC-MAC utilize r to achieve a diversity gain of 2. Although in the case of DRRS-MAC diversity gain of 2 gets a chance only if transmission via direct link fails.

Contention based MAC are more commonly seen than reservation based. Most of the protocols modify RTS/CTS/ACK control packets as they deem fit like DRRS-MAC, CoopMC-MAC, CoopMAC, VC-MAC, except for IrcMAC protocol which does not use them all together. Variations include HTS/RTS/CTS in CoopMAC or CTU/CTU-R/CTU-ACK in DRRS-MAC etc.

Number of relay nodes used varies from 1 to many. The question here arises: if only one relay is enough than what is the need for a large number of relay nodes to be tested and used as in CoopMC-MAC. While one cannot ignore the fact that one relay does not leave much choice for optimal relay selection, it also has to be kept in mind that a large pool of relay nodes to choose from can also cause wastage of energy and time for source. Moreover the source also has to choose a good metric for measurement of optimality of relay in case of a large number of relay nodes. Therefore how many number of r to be used is a very good path for research.

For the protocols which use only one relay, it is assumed that relay used is the optimal one. Another question to be asked here is; is this optimality of the relay measured with respect to the source or the destination or both? If this question goes unanswered, the results can pose a large threat to the output BER.

The overhead due to control packet frames is incurred most in CoopMC-MAC as the protocol allows the use of a large number of control packet frames like CTU/CTU-R/CTU-ACK packets. Next in line are DRRS-MAC and C-TDMA protocols. The main idea is to reduce overhead incurred due to control packets frames as much as possible. This requires that less and less number of control packets frames be used. For example in IrcMAC the control packets used is only restricted to RTS/CTS/ACK. This not only uses less energy but also reduces delay in the channel data transmission. It will be interesting for the reader to note that most of these control packets have the same frame structure as the basic RTS/CTS/ACK control packets of IEEE802.11

protocol. This again implies that instead of using multiple copies/types of the same control packet frames, the same ones can be used. Hence they issues pose very good research areas.

The metrics used in the protocols described here vary, including i) distance and SNR, ii) time taken for destination to send CTS, iii) placement of nodes around the source-destination overlapping area, and iv) reception of CTS/HTS packet from relay. Measurement of distance may be considered quite a good metric. Nonetheless the best metric is the one which does not cause redundant delay.

Furthermore, it is seen in most of these protocols that the time taken for transmission was not considered important. In protocols like CoopMAC-MAC, the entire overhead must take a lot of time but it was not given much thought. This can cause delay and propositions for its reduction of imply a very good research area.

Albeit the harsh analysis, it can be said that IrcMAC proposes a great protocol including the RR frame and relay selection. But it falls short when addressing the issue of how source will store such a large amount of data related to a large number of relay nodes, which according to the paper total a figure of 500. It fails to mention what will happen if the transmission link from source to destination is very large and full of fading, and there is no relay nodes to fill in that gap because of transmission times more than coherence time of the channel, in which case they will not be used.

All protocols talk at length about metrics concerning distance, SNR or time for transmission, or diversity gain of 2. Hence they forget an important issue; redundancy of number of relay nodes used. This can be explained by considering an optimal relay and source. If the use of either gives the same result in terms of SNR, BER, saturated throughput, energy efficiency and network density, etc. then instead of using both, will it be better to use 1? Using both relay node and source can give a diversity gain of 2 but at what cost? This also allows for a vast amount of research to answer these questions.

None of these protocols have been applied in a practical setting, so furthering of the research in a physical environment can be a great field for modelling and simulation.

A robust Cooperative MAC protocol is the one which successfully mitigates fading but also does not waste too much energy. It should be good at avoiding

collision. A relay based system is supposed to help a wireless network deliver data packets in adverse conditions without adding to the adversity. Therefore devising protocols which do not offer an efficient solution to these conditions are just pointless. It is also seen that when cooperative diversity is implemented in real-life situation, several challenges are faced; modelling Rayleigh multipath fading channels instead of Rayleigh flat fading or AWGN channels creates a very difficult scenario and in both contention based or reservation based relay selection is very difficult keeping in mind the requirement of the system.

If a single optimal relay is to be chosen from a selection of relay nodes, then it should be noted that the node chosen fulfils the following options:

- Relay selection should ensure low collision probability
- It should not introduce redundancy but should also try to increase the diversity gain.
- Optimal relay selection should be time efficient;
- There should be a definite guarantee that the relay selected is the best one.
- It should also be noted that relay chosen should suit the needs of the particular wireless environment.

The last point can be explained by an example. If the situation consists of mobile nodes then it is highly unfeasible that source first tries to send data to destination with having a relay chosen beforehand because that relay may not be available for cooperation due to its mobile nature, in case the data deliverance to the destination fails.

Coming onto distributed networks, every node has the right to send a packet and each can overhear the other nodes. Therefore in case of a system where the data transmission and/or cooperation is not controlled by any nodes (particularly source) an important question left unanswered is; how do all the nodes know which node has to forward a data packet in case of failure in delivery? The answer to this question can possibly result in a decentralized distributed system and presents a large research area.

One of the best proclaimed works was done by CoopMC-MAC on multi-channels which are a good cooperative MAC protocol technique. They are good

at mitigating channel interference and its effects. Therefore they can enhance the performance of wireless networks.

A comparison of all the listed protocols is given in Table 2. The columns represent Backward compatibility with IEEE802.11 protocols, which is given in most cases. Third column gives the number of relay nodes which are used for cooperation whereas fourth column represents the type of cooperation. The types that are most commonly seem include cooperative diversity (antenna diversity), cooperative MIMO and cooperative relaying. Fifth column explores the type of MAC protocol used and sixth gives the assumed position of potential relay nodes which in most cases is given to be between source and destination except for DRRS-MC which gives a truer picture of the placement of relay nodes. Cooperation decision, i.e. whether the source makes the decision to cooperate and chooses a relay node or each relay node is independent to present themselves as optimal relays is shown in seventh column. Whether overhead due to control packet frames is incurred or not is indicated by eighth column while ninth column concludes the table by listing the value of diversity gain for each protocol.

IV. CONCLUSION

It is a smart idea to incorporate cooperative communication in today's wireless networks. This paper focuses on the different types of MAC protocols incorporating cooperative communications in them to form a better network, eliminate fading, reducing transmission and reception energy and causing an increase in diversity. The techniques explained are all implemented in ad-hoc networks. Each protocol is unique in its own stance and takes advantage of either the spatial reuse or control packet reuse or distributed nature etc. Thus each protocol revolutionises a method for cooperative communications. In the end the methods are all compared and a very comprehensive analysis is given. It is hard to say which protocol is better than others or which is the best. Suffice it to say that the direction of the intended research defines which protocol to be preferred.

This paper also aims to make it easier for the upcoming researchers who wish to undertake research in this field of wireless networks to have a quick and fast, but comprehensive view on the current unanswered issues of the network, further giving ideas where appropriate for further research.

Table 2: Comparison table for different Cooperative MAC protocols

| Cooperative MAC Protocol | Backward Compatibility | Number of relay nodes used for cooperation | Type of cooperation | MAC protocol | Assumed position of potential relay | Cooperation decision | Over-head | Diversity gain |
|---------------------------------|------------------------|--|---------------------------------------|--------------|-------------------------------------|----------------------|-----------|----------------------------|
| CoopMAC | Yes | 1 | Cooperative Relaying | Contention | Between source and destination | Source-dependent | Yes | 1 |
| CD-MAC | Yes | 2 | Cooperative Relaying | Contention | Between source and destination | Source-dependent | Yes | 2 or more |
| C-MAC | Yes | 2 | Cooperative Relaying | Contention | Between source and destination | Source-dependent | Yes | 2 or more |
| C-TDMA | Yes | All available relays | Cooperative Diversity | Reservation | Between source and destination | Source-dependent | Yes | N number of diversity gain |
| IrcMAC | Yes | 1 | Cooperative Relaying | Contention | Between source and destination | Source-dependent | Yes | 1 |
| VC-MAC | Yes | 2 | Cooperative MIMO | Contention | Between source and destination | Source-dependent | No | 2 or more |
| CoopMC-MAC | Yes | 2 or more | Multiple channel Cooperative Relaying | Contention | Between source and destination | Source-dependent | No | N number of diversity gain |
| (DRRS) Cooperative MAC protocol | Yes | 1 | Cooperative Relaying | Contention | Anywhere in the network | Source-dependent | Yes | 1 |

REFERENCES

- [1] Daquan Feng, Chenzi Jiang, Gubong Lim, Leonard J. Cimini, Jr., Gang Feng, Geoffrey Ye Li, "A Survey of Energy-Efficient Wireless Communications", IEEE Communications Surveys & Tutorials, Vol. 15, No.1
- [2] Eun-Sun Jung, Vaidya, N.F., "An energy efficient MAC protocol for wireless LANs", INFOCOM 2002, Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. (Volume:3), Page(s): 1756 - 1764 vol.3
- [3] Unnikrishnan M1, Rinu Titus2, Premkumar C V, "Mac Protocols For Cooperative Diversity In Wlan", International Journal of Research in Engineering and Technology, Volume: 03 Special Issue: 15, Dec-2014.
- [4] Jun Peng, Liang Cheng, Sikdar, B., "A Wireless MAC Protocol with Collision Detection", Mobile Computing, IEEE Transactions on (Volume:6 , Issue: 12), Page(s): 1357 – 1369.
- [5] D Saha, M R Yousuf, M A Matin, "ENERGY EFFICIENT SCHEDULING ALGORITHM FOR S-MAC PROTOCOL IN WIRELESS SENSOR NETWORK ", International

- Journal of Wireless & Mobile Networks (IJWMN) Vol. 3, No. 6, December 2011
- [6] Mthulisi Velepini, Mqhele E. Dlodlo, "The analysis of multichannel MAC protocols which implement a control channel", International Journal on AdHoc Networking Systems (IJANS) Vol. 4, No. 1, January
- [7] Shams ur Rahman, Mushtaq Ahmad, Shafaat A. Bazaz, "A new Energy-Efficient TDMA-based MAC Protocol for Periodic Sensing Applications in Wireless Sensor Networks", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 1, July 2012
- [8] Jamila Bhar, "A Mac Protocol Implementation for Wireless Sensor Network", Journal of Computer Networks and Communications, Volume 2015 (2015), Article ID 697153, 12 pages
- [9] Prakash Ranganathan, Kendall Nygard, "Time Synchronization In Wireless Sensor Networks: A Survey", Internayional Journal of UbiComp, Vol 2, April 2010.
- [10] IEEE Standard 802-1990: "IEEE Standards for Local and Metropolitan Area Networks: Overview and Architecture,"
- [11] Ernesto Zimmermann, Patrick Herhold, Gerhard Fettweis, "On the Performance of Cooperative Relaying Protocols in Wireless Networks", Transactions on Emerging Telecommunications Technologies, Volume 16, Issue 1, pages 5-16, January/February 2005
- [12] Weifeng Su, Xin Liu, "On Optimum Selection Relaying Protocols in Cooperative Wireless Networks", IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. 58, NO. 1, JANUARY 2010
- [13] Nandar Lynna, Osamu Takyu, Koichi Adachi, Masao Nakagawa, Tomoaki Ohtsuki, "An Autonomous Selective Cooperative ARQ Protocol for Hybrid Mobile Wireless Sensor Networks", JOURNAL OF COMMUNICATIONS, VOL. 6, NO. 2, APRIL 2011
- [14] Vishal K. Shah, Anuradha P. Gharge, "A Review on Relay Selection Techniques in Cooperative Communication", International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 5, November 2012
- [15] Pablo Brenner, "A technical tutorial on the IEEE 802.11 protocol", BreezeCom Wireless Communications 01/1997.
- [16] Pei Liu, Zhifeng Tao, Sathya Narayanan, Thanasis Korakis, Shivendra Panwar, "CoopMAC: A Cooperative MAC for Wireless LANs", IEEE Journal on Selected Areas in Communication (JSAC), VOL. 25, NO. 2, FEBRUARY 2007.
- [17] Brandwein, R., Cox, T., Dahl, J, "The IEEE 802.6 physical layer convergence procedures", LCS, IEEE, May 1990, Volume:1, Issue: 2, Page(s): 26 - 34
- [18] Sangman Moh, Chansu Yu, Seung-Min Park, and Heung-Nam Kim, "CD-MAC: Cooperative Diversity MAC for Robust Communication in Wireless Ad Hoc Networks", Communications, 2007. ICC '07. IEEE International Conference on, 24-28 June 2007 Page(s):3636 - 3641
- [19] Aytac Azgin, Yucel Altunbasak, Ghassan AlRegib, "Cooperative MAC and Routing Protocols for Wireless Ad Hoc Networks", Conference: Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE, Volume: 5.
- [20] Zhuo Yang, Yu-Dong Yao, Xiaochen Li, and Di Zheng, "A TDMA-Based MAC Protocol with Cooperative Diversity", IEEE COMMUNICATIONS LETTERS, VOL. 14, NO. 6, JUNE 2010.
- [21] Murad Khalid, Yufeng Wang, Ismail Butun, Hyung-jin Kim, In-ho Ra and Ravi Sankar, "Coherence time-based cooperative MAC protocol for wireless ad hoc networks", Journal on Wireless Communications and Networking 2011, 2011:3.
- [22] Jin Zhang, Qian Zhang, Weijia Jia, "VC-MAC: A Cooperative MAC Protocol in Vehicular Networks", IEEE Transactions On Vehicular Technology, Vol. 58, No. 3, March 2009.
- [23] Devu Manikantan Shila, Tricha Anjali, Yu Cheng, A Cooperative Multi-Channel MAC Protocol for Wireless Networks", Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE, Page(s): 1 - 5
- [24] Md. Rasel Rana, Md. Faruque Hossain, Md. Kamrul Hasan, Md. Uzzal Hossain, Md. Obaidur Rahman, "Distance-Aware Reliable Relay Selection Provisioning in Wireless Cooperative MAC Protocol", Strategic Technology (IFOST), 2014 9th International Forum on, Page(s):156 - 160

Using Cohesion and Capability Maturity Model Integration (CMMI) as Software Product and Process Quality criteria: a case study in Software Engineering practice in Botswana

Ezekiel U Okike and Motsomi Rapoo

Computer Science Department, University of Botswana
Gaborone, 00000, Botswana

Abstract

The need to ensure quality in software engineering practice necessitated the introduction of software measurement and other quality standards introduced by the software engineering institute. The appropriate introduction of quality standards however does not assume that practitioners easily adapt to its utilization without assistance by experienced professional. This study aims to investigate software engineering practice in Botswana in order to assist software companies understand and use software engineering quality measures and standards. 14 software companies were identified out of which 7 indicated interest in participating in this study. The result indicates that most of the participating companies are yet to satisfy the Capability Maturity Model Integration (CMMI) 18 key performance areas at the 5 levels. Of the 5 companies which indicated that they are using CMM/CMMI standard, only 1 company satisfies 100% CMMI requirements. The study reveals the need to launch a programme to bring CMMI into Botswana and to train software companies on the use of appropriate software metric tools in order to ensure software quality in Botswana

Keywords: *Software engineering, software measurement, software quality, cohesion, capability maturity model integration*

1. Introduction

Software Engineering (SE) is concerned with all aspects of software production from the early stages of system specification through to maintaining the system after it has gone into use ((Sommerville, 2011).). It describes the collection of techniques that apply an engineering approach to the construction and support of software products. The engineering approach of SE implies that that activities are well understood and controlled, so that there are few surprises as the software is specified, designed, built and maintained. According to Fenton & Pfleeger (1997), “while Computer Science (CS) provides the theoretical foundations for building software, software engineering focuses on implementing the software in a controlled and scientific way” so that the resulting software is reliable, and acceptable because it is of the right quality. Like other engineering disciplines, SE applies theories, methods, models and tools which ensures that a rigorous approach is followed in developing the software. In addition to addressing technical processes of software development, SE include activities such as software project

management, development of tools, methods and theories to support the production of quality software. Two main approaches to ensure software quality may be described in terms of the application of software metric/measurement and enforcement of software improvement standards such as the Capability Maturity Model Integration (CMMI), ISO 9001, ISO9002, ISO9003, SPICE, IDEAL etc)

Measurement in Software Engineering (SE) is necessary to show that Software Engineering also go through required formalism and rigorous approaches like other well established science and engineering disciplines. According to Fenton and Pfledger, 1997, measurement may be described as a mapping from the empirical world to the formal, relational world. Consequently, a measure is the number or symbol assigned to an entity by this mapping in order to characterize an attribute. Two most important software quality measures are cohesion and coupling (Chidamber and Kemerer, 1996). While cohesion refers to the degree of relatedness among the members of a class (or a module), coupling deals with the interconnectedness of modules in a program. Practically, high cohesion and low coupling implies good software design. High cohesion, range [0,1] indicates good design and good quality software (Okike and Osofisan, 2008; Okike, 2010, Okike, 2007, Chidamber and Kemerer, 1994)

In this, we describe the use and application of cohesion metric in ensuring software quality. Additionally, we investigate into software engineering practice in Botswana with a view to ascertaining if quality procedure and standards are followed by the local developers

2. Problem statement

Software processes and practice in Botswana have not been fully investigated. One study by Ayalew and Motlhala (2014) suggest that software companies are reluctant to participate in software process improvement studies which could provide useful insight to the state of software practice in the country. As the government through the appropriate organ is currently pursuing a vibrant Information and communication Technology (ICT) policy, they have identified software industry as one of the strategic priority areas in the policy. Therefore, to provide input into the ICT strategic policy, it is necessary to investigate the state of software practice in the country. This will also ensure that appropriate standards are

followed in software development efforts so that we can have quality software developed locally

2. Aims and objectives

The main objective of this study is to investigate software engineering practice in Botswana in order to ascertain if software process and process improvement standards are followed by the software companies. In particular the capability and maturity of software companies in Botswana are investigated

3. Methodology

Fourteen (14) Software companies in Gaborone and Lobatse were initially selected from the list of government approved ICT companies in Botswana. A pre test questionnaire was administered on the companies focusing on use of software improvement standards and models in Botswana. The questionnaire was improved and 7 companies which indicated having used/still using standards such as the Capability Maturity Model (CMM), Capability Maturity Model Integration (CMMI), ISO 9001, ISO9002, ISO9003, SPICE, and IDEAL were fully investigated on their capability and maturity levels on the CMMI 18 key performance areas (KPA's). The rest of this paper counts of 7 sections. Section 2 looks at some important literatures with regards to software quality measures and metrics as well as models. Section 3 explains the methodology of this paper. Section 4 explains the mathematical formalism behind the class cohesion metric- the Lack of cohesion in methods (LCOM). Section 5 investigates software practice in Botswana and the capability maturity level of participating companies on the CMMI scale. Section 6 is conclusion and section 7 is references.

4. Literature review

Software quality is an essential part of software process activities, products and resources. Fenton and Pfledger (1997) identified processes, products and resources as essential software entities and attributes that must be measured to assure quality . While processes are collections of software related activities; products are artefacts, deliverables or documents resulting from process activities and resources are entities required by a process activity (ibid). In addition, software artefacts have two types of attributes namely internal attributes and external attributes. The internal quality attributes of a software are measured in terms of the product itself and these include module cohesion, module coupling, data structure, algorithm, nesting level, and control structure (Okike, 2007). These attributes have been identified as the sources of structural complexity in software and if proper measures

are applied such complexities could be reduced resulting to high quality software (Darcy et al, 2005; Hartmanis, 1994; Gorla and Rakrishnan, 1997; Page-Jone, 1998). Furthermore, of the seven identified sources of software complexity, cohesion and coupling are said to be the dominant attributes (Darcy, Kermerer, Slaughter and Tomayko ,2005). Hence by using appropriate cohesion metric and coupling tools, software quality can be improved at both design and code levels (Okike, 2007; Okike and Osofisan, 2008; Okike, 2010; Bieman and Kang, 1998; Bieman and Kang, 1995, Bieman and Ott, 1994

Apart from the use and application of software metric and measurement tools, software standards and processes in use can also affect software quality. For example understanding and applying the Capability Maturity Model Integration (CMMI) suite can improve software process and hence the quality of process activities (Dangle, Larsen, Shaw and Zelkowitz, 2005; JCSE, 2008). The University of Johannesburg Centre for Software Engineering launched a programme to introduce CMMI in South Africa in 2006 by training local resources to be able to offer CMMI consulting in order to grow the country's capacity to deliver world class software by 2030 (JCSE, 2008). The initiative has paid off as CMMI has been widely accepted and being used in South Africa by most companies. In this study, our focus is on the application of a cohesion metric tool (for example the Lack of Cohesion in Methods (LCOM) tool) to assure product quality and the application of CMMI to assure process quality in software engineering by software companies in Botswana

5. Measuring Functional cohesion

Different researchers defined cohesion measures in the procedural paradigms. Examples include Bieman & Ott (1994), Bieman & Kang (1998). Also several measures exist in the Object Oriented paradigm for measuring class cohesion (Okike, 2007). For example Chidamber and Kermerer's Lack of Cohesion in Methods (LCOM) metric (Chidamber and Kermerer, 1994)

5.1 The Lack of cohesion in methods metric (LCOM)

Consider a class C1 with n methods M1, M2, ..., Mn. Let {Ii} = set of instance variables used by method Mi. There are n such sets {I1}, ..., {In}. Let P = { (Ii, Ij) | Ii ∩ Ij = ∅ } and Q = { (Ii, Ij) | Ii ∩ Ij ≠ ∅ }. If all n sets { I1 }, ..., {In} are ∅ then let P = ∅

$$\text{LCOM} = \{ |P| - |Q|, \text{ if } |P| > |Q| \\ = 0, \text{ otherwise}$$

Example: Consider a class C with three methods M1, M2 and M3. Let {I1} = {a,b,c,d,e} and {I2} = {a,b,e} and {I3}

= {x,y,z}. {I1} ∩ {I2} is nonempty, but {I1} ∩ {I3} and {I2} ∩ {I3} are null sets. LCOM is (the number of null intersections – number of non empty intersections), which in this case is 1.

The theoretical basis of LCOM uses the notion of degree of similarity of methods. The degree of similarity of two methods M1 and M2 in class C1 is given by:

$$\sigma(\) = \{I1\} \cap \{I2\}$$

where {I1} and {I2} are sets of instance variables used by M1 and M2 . The LCOM is a count of the number of method pairs whose similarity is 0 (i.e. σ() is a null set) minus the count of method pairs whose similarity is not zero. The larger the number of similar methods, the more cohesive the class, which is consistent with the traditional notions of cohesion that measure the inter relatedness between portions of a program. If none of the methods of a class display any instance behaviour, i.e. do not use any instance variables, they have no similarity and the LCOM value for the class will be zero.

The LCOM value provides a measure of the relative disparate nature of methods in the class. A smaller number of disjoint pairs (elements of set P) implies greater similarity of methods. LCOM is intimately tied to the instance variables and methods of a class, therefore is a measure of the attributes of an object class.

5.1.1 Definition 2.

As explained by Hitz and Montazeri(1996) and also in Okike (2010a), Okike (2010b)

$$\text{Let } P = \emptyset, \text{ if } \text{AR}(m) = \emptyset \forall m \in \text{MI}(c) \\ = \{ \{m1, m2\} \mid m1, m2 \in \text{MI}(c) \wedge m1 \neq m2 \wedge \text{AR}(m1) \\ \cap \text{AR}(m2) \cap \text{AI}(c) = \emptyset \}, \text{ else}$$

$$\text{Let } Q = \{ \{ m1, m2 \} \mid m1, m2 \in \text{MI}(c) \wedge m1 \neq m2 \wedge \\ \text{AR}(m1) \cap \text{AR}(m2) \cap \text{AI}(c) \neq \emptyset \}$$

$$\text{Then } \text{LCOM2}(c) = \{ |P| - |Q|, \text{ if } |P| > |Q| \\ = 0, \text{ otherwise}$$

This is a refinement of previous definition to include inherited methods and attributes

LCOM2 of many classes are set to zero although different cohesions are expected

$$\text{Let } P = (n - 1) \tag{1}$$

$$Q = n-1 \quad (2)$$

So that LCOM

$$P-Q = \frac{n!}{(n-2)!2!} - 2(n-1) \quad (3)$$

From (3), for $n < 5$, $LCOM = 0$;
for $n \geq 5$, $1 < LCOM < n$

5.2 Empirical applications

Measuring class cohesion have been used in many empirical studies to show properly designed software (Okike, 2007; Badri & Badri, 2004; Shumway, 1997; Amman, Yamasski and Nodi, 2002). More recent application of the cohesion metric include the following:

- Measuring Class cohesion in OO systems from 6 industrial systems (over 3000 classes) (Okike, 2007)
- Measuring students problem solving and decision making ability based on Personality traits (Okike and Amoo, 2014)
- Assessment of Code level quality of Programs by different programmers (Okike, 2014)

Although there are different levels of cohesion such as coincidental, logical, temporal, procedural, communicational, sequential and functional, however, functional cohesion is the most desirable because it performs exactly one action or achieves a single goal (Page-Jone, 1988; Okike, 2007; Okike, 2010a). Such a module is highly reusable, relatively easy to understand (because you know what it does) and is maintainable.

6.1 Software Engineering practice in Botswana

Reports from our current research efforts is suggesting that Botswana might face software Crisis if urgent quality measures are not put in place (Ayalew and Motwala, 2014; Okike and Rapo, 2015). Software Companies in Botswana do not want to participate in CMM based research, suggesting that they are not following the Capability Maturity Model (CMM) standard

6.1 Using Capability maturity model Integration to ensure software quality

The Capability Maturity Model Integration (CMMI) measures the capability of organizations based on 18 key performance areas (KPAs) at 5 levels (Dangle, Larsen and Shaw, 2005) :

- Level 1 Initial
- Level 2 Repeatable Consist of 6 KPAs

- Requirements Management
- Software Project Planning
- Project tracking and oversight
- Subcontract Management
- Quality management
- Configuration management
- Level 3 Defined, consist of 7KPAs
 - Organization Process focus
 - Organization Process definition
 - Training Program
 - Integrated software Management
 - Software product engineering
 - Intergroup coordination
 - Peer reviews
- Level 4 Managed 2KPA
 - Quantitative Process Management
 - Software Quality Management
- Level 5 Optimized , Consist of all 18 KPAs including

Defect prevention

Organizations are rated using this KPAs for success and maturity, and also ability to handle projects. An organization that succeeds at none of the KPAs is rated at the lowest level. In this study, 14 Software organizations participated and 7 indicated certification with CMM/CMMI. We selected the 7 companies and measured their CMMI level based on the CMM 18 key performance areas (KPAs) (Rapo and Okike, 2015).

6.3 Results and Discussion

The findings of the study are presented as shown in figures 1-4 below. Company identifiers appear on the vertical axis while CMMI percentage coverage by the companies appear on the horizontal axis

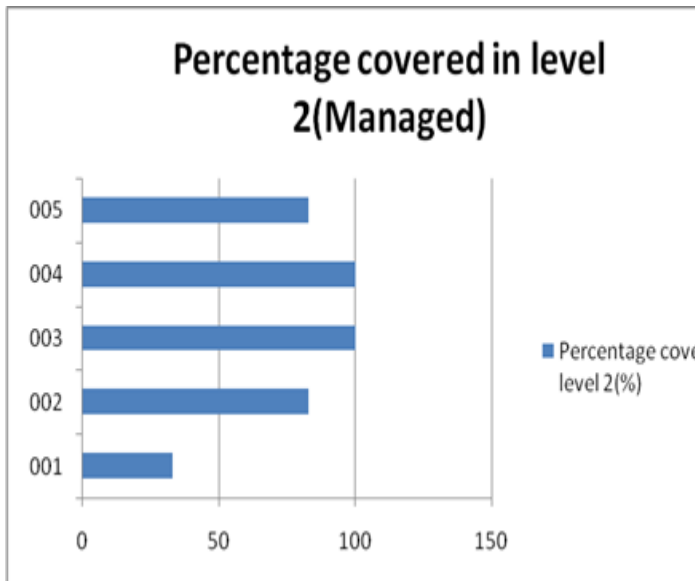


Figure 1. measuring CMMI at level 2 for companies 001-005

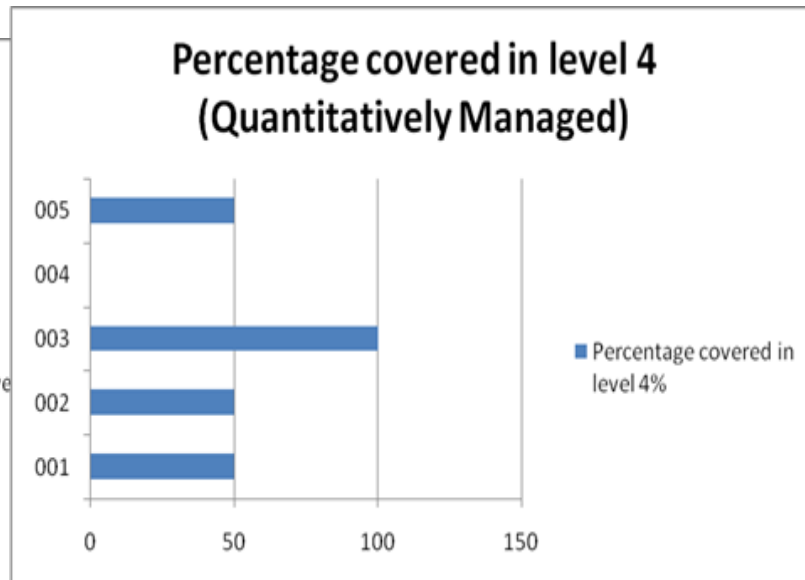


Figure 3. measuring CMMI at level 4 for companies 001-005

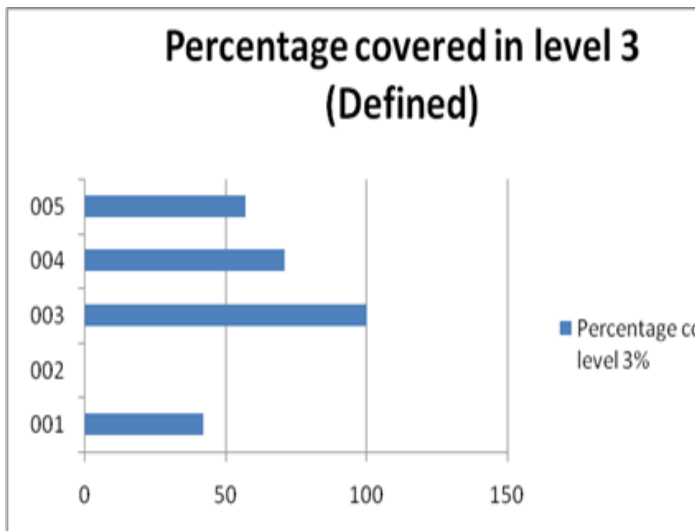


Figure 2. measuring CMMI at level 3 for companies 001-005

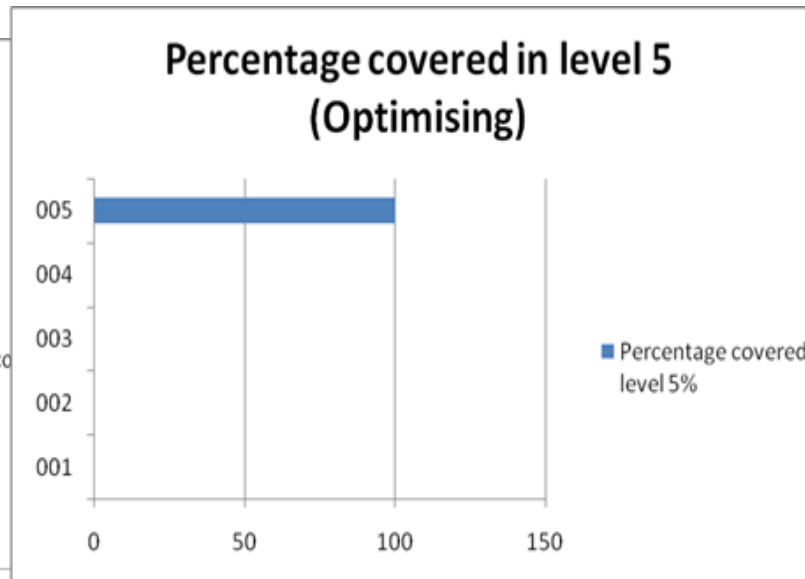


Figure 4. measuring CMMI at level 5 for companies 001-005

From figure 1, companies 003 and 004 satisfied CMMI requirements at 100% of level 2 on the CMMI scale. Companies 002 and 005 satisfied about 80% requirement of CMMI at level 2, while company 001 did not get up to 50% on the CMMI scale.

From figure 2, only company 003 met 100% CMMI requirements at level 3, company 004 met about 75%, while company 005 met about 60%. Company 001 did not meet up to 50% of the CMMI requirements at level 3.

From figure 3, only company 003 met 100% CMMI requirements at level 4, the other companies 002, 004 and

005 met 50% CMMI requirements at level 4. Company 001's claim of meeting 50% CMMI requirements at level 4 is doubtful since the company did not meet 50% requirements of CMMI levels 2 and 3.

From figure 4, company 005 claim of meeting 100% CMMI requirements at level 5 is doubtful since the company did not meet CMMI requirements at levels 2, 3 and 4.

6.4 Discussion

From this and other studies of software practice in Botswana, it is clear that the software companies are not meeting CMM/CMMI quality assurance requirements based on the 18 key performance areas. Hence, the need for Botswana to launch a programme to bring CMMI into mpromised. Available metric tools as well as software improvement models should be utilized effectively in software production so as to assure quality in both software process activities and the product itself. Botswana's ICT policy has the vision of making the country a "globally software practices in the country is one of the ways in achieving this noble ICT vison. Therefore the application and utilization of appropriate technologies and models by software professionals in the country is unavoidable. As such, relevant research institutions and development for the benefit of the country

7 References

- [1] Aman, H; Yamasski, K; Noda, M (2002). A proposal of class cohesion metrics usingsizes of cohesive parts, *Knowledge based Software Engineering*. T Welzer et al. eds. IOS press, pp 102-107
- [2] Ayalew, Y and Motlhala, K (2014) Software process practice in small software companies in Botswana. Computational Science and Its Applications (ICCSA), 2014 14th International Conference on June 30 2014-July 3 2014
- [3] Badri, L and Badri, M (2004). A proposal of a new class cohesion criterion: an empirical study. *Journal of Object Technology* 3, 4:145-159
- [4] Bieman, J. M and Kang, B. K 1998. Measuring Design –level Cohesion. *IEEE Transactions on Software Engineering*, vol. 20, no. 2, pp. 111-124.

the country in line with what South Africa did in 2006 (JCSE, 2008). In this direction, the Software Engineering (SE) research Group at University of Botswana plans to work with appropriate government agency to introduce CMMI use in Botswana. Some pilot companies could be selected and initially trained on the use of CMMI to assist them to implement CMMI properly and get them recognized officially as capable and mature software companies in Botswana. Companies that participated in this study are likely pilot candidates.

6.5 Conclusion

The need for Quality in all aspects of Software Engineering (SE) can not be Co competitive, knowledge and information society where lasting improvements in social, economic and cultural development is achieved through effective use of ICT" (New National Policy on ICT, 2004). Adherence to quality procedures in government agencies should partner in the drive for quality in Software practice and product in the country as is the case in South Africa and other countries. The partnership should bring about a coordinated effort in ICT research and

- [5] Bieman, J. M and Kang, B. K 1995. Cohesion and reuse in Object oriented system. Proceedings of the symposium on software reusability (SSR '95), Seattle: WA. Pp. 259-262
- [6] Bieman, J. M and Ott, L. M. "Measuring Functional Cohesion", *IEEE Transactions on Software Engineering*, vol. 20, no. 8, pp. 644-658.
- [7] Dangle, K. C; Larsen, P and Shaw, M (2005). Software Process Improvement in organizations: a case study. *IEEE Software*, Nov/Dec. pp 68-75
- [8] Darcy, P. D; Kermerer, C. F; Slaughter, S.A and Tomayko, J.E (2005). The structural complexity of software: An experimental test. *IEEE Transactions on software Engineering*, 32, 1:54-64
- [9] Fenton, N and Pflugger, S (1998) *Software Metrics: A Rigorous and Practical Approach*, 2nd ed. Boston, MA:PSW Publishing
- [10] Gorla, N and Ramakrishnan, R (1997). Effect of software structure attributes on software

- development productivity. *Journal of systems and software* 36,2:191-199
- [11] M. Hitz and B. Montazeri(1996). Chidamber and Kenmerer 's mtric Suite: A Measurement Theory Perspective, *IEEE Transactions on Software Engineering*, vol. 22. no. 4, pp.267-270
- [12] Okike, E. U (2007). Measuring class cohesion in Object Oriented systems using Java and Chidember and Kemera Java metric . PhD Thesis, Department of Computer Science, University of Ibadan, Unpublished.
- [13] Okike, E and Osofisan, A.(2008). An evaluation of Chidamber and Kermerer's Lack of cohesion in methods metric using different normalization approaches. *Afr. J. comp. & ICT*, Vol. 1, No 2. Pp. 35-
- [14] Okike, E (2010a), "A pedagogical Evaluation and Discussion about the Lack of Cohesion in methods (LCOM) Metric Using field Experiment", *International journal of computer science issues*, Vol. 7, issue 2, No. 3, March 2010. Pp 36-43
- [15] Okike, E U (2010b), A proposal for a normalized Lack of cohesion in methods metric using field experiments. *IJCSI International Journal of Computer Science Issues*, Vol. 7, Issue 4, pp.19-27
- [16] Okike, E. U and Amoo, A. A (2014). Problem Solving and Decision Making: Consideration of Individual Differences in Computer Programming Skills Using Myers Briggs Type Indicator (MBTI) and Chidamber and Kemerer Java Metrics (CKJM). *Journal of Applied Information Science and Technology*, 7:1 pp 27-34
- [17]Okike, E U (2014). A code level based programmer assessment and selection criterion using metric tools. *(IJACSA) International Journal of Advanced Computer Science and Applications*, Vol. 5, No. 11, pp 85-91
- [18] Shumway, M. F (1997). Measuring class cohesion in Java. Master of Science Thesis, Department of Computer Science, Colorado state University, Technical Report CD-97-113
- [19] Page-Jones, M (1988). The practical guide to structural system design. 2nd ed. Englewood Ciff, New Jersey: Yourdon press, prentice Hall
- [20] Rapo, M (2015). Towards a Software Process Selection Improvement (SPI) for small and medium companies in Botswan. MSc , Department of Computer Science, University of Botswana Dissertation. Unpublished
- [21] Sommerville, I (2011) *Software Engineering*, 9th Edition Addison Wesley:NY.
- [22] Wits University, Joburg Centre for Software Engineering, JCSE (2008). South Africa moves towards process improvement. Press release. http://www.itweb.co.za/index.php?option=com_content&view=art..



Dr Ezekiel Uzor Okike, MIEEE, MACM, MAIS is a Senior Lecturer in the Department of Computer Science, University of Botswana, Gaborone. He holds a PhD in Computer Science (2007), a Master of Information Science (M.Inf,Sc), 1994 and a B.Sc Computer Science, (1992) all from the University of Ibadan, Nigeria. He is a member of IEEE, ACM, and AIS. Presently he is the cluster chair of the Information Systems Cluster, Department of Computer Science, University of Botswana. He has published many papers in international journals and attended many international conferences where he presented papers. His research interests are in Software Engineering, Information Systems, Cyber Security, Programming languages, Compilers, and High Performance computing



Mr Motsoni Rapo is a Post graduate student in the Department of Computer Science, University of Botswana. He holds a Bachelor of Science in computer science, and recently completed the Master of Science (M.Sc.) degree in Computer Science from the University of Botswana. His research interest is in the area of Software Engineering.

Trust: Models and Architecture in Cloud Computing

Usvir Kaur

Research Scholar,
I.K.Gujral Punjab Technical university,
Jalandhar

Dheerendra Singh

Professor,
Deptt. of Computer Science & Engineering
S.U.S. College of Engg. and Technology,
Tangori, Mohali, Punjab, India.

Abstract: In today era, Cloud Computing has become an emerging technology. Various service providers are available in market with different configurations. It becomes very difficult for users to choose between various cloud service providers. Trust plays a vital role in field of cloud computing services such that it enables users to choose particular cloud service provider from available service on basis of how this technology behaves in past. This paper discusses various parameters and trust model framework. Different trust models for single web service and various parameters used by them for calculating trust are reviewed.

Keywords: *cloud computing, trust model architecture, security.*

I. INTRODUCTION

Cloud computing is a platform, which relies on virtualization technology for sharing resources and it provides elastic, on-demand and instant services to its customers. On the basis of usage it charges bill form customers known as utility bill. Service ,Basically , after addition of some characteristics such as resource sharing using virtualization, elastic, on demand and instant service data centre becomes cloud computing. We may call cloud computing word after data centre as providers. According to the law of nature, any change in some infrastructure where certain benefits are and risks, cloud computing is not an exception. Due to the some characteristics, such as shared, on demand nature of cloud computing [1]. Cloud computing becomes vulnerable to some risks.

Cloud computing evolves from various relevant technologies such as grid computing, cluster computing, utility computing, and distributed computing. There are several security issues that are inherited by cloud computing. Leavitt [1] describes six challenges which are: control; performance, latency and reliability, security and privacy, related bandwidth costs, vendor lock-in and standards, and transparency. Cloud computing should not be compared with resource provisioning infrastructure. It is different from resource provisioning in following ways. [2] Cloud Computing environment infinitely scalable and it provides platform for various applications e.g. Social networking sites such as face-book, twitter. Moreover cloud platform can be used for every purpose such as disaster recovery, business continuity to fully outsourced internet and communications technologies service. Cloud providers rely data centres for outsourcing. So it becomes important issue of security and trust when a cloud client places it valuable resources such as knowledge, information and data on external service. [2]

There are various threats involved in cloud computing related to security of cloud servers. Basically, security of data centers and computer networks is still a challenging task, moreover characteristics such as shared, on-demand nature makes it more difficult.

Virtualization is a key factor in success of cloud computing because it helps in multiplexing resources on a single platform for multiple users. So it becomes an essential issue for monitoring the behavior of virtual machines (VMs).

II. NEED OF TRUST

According to the Gartner analysis CLOUD computing reached a peak of inflated expectations in 2009 / 2010 and is now in the phase of increasing disillusionment where users become more and more aware of the deficiencies of the system and have to start assessing whether and what is worth moving to the CLOUD.[3] Trust is outcome of various parameters which are incorporated such as identity management, authorization, authentication etc.

Trust covers all aspects of security within single term. [2]

Various Parameters

- a) Identity Management System: - Identity management is a primary element for the security of cloud system and for every internet applications. Security strength can be measured by generating identities for every client by service providers. Identity Management process includes various parameters such as identity creation, storage and life cycle management of the identity. These parameters can be measured against IDM strength component of trust model.
- b) Authorization:-this trust component checks various actions performed by the users so that user should be restricted to perform any action which is not meant for. This authorization action could include service access, performing any operation input/output related activities. This method is measured with respect to authorization strength.
- c) Authentication: - an authentication check would increase user confidence at time of login and identity verification process. Basically, it plays a double role by giving privilege to both parties client and provider to trust each other. Authentic use of cloud service by legitimate provider can determined strength of authentication.
- d) Data Protection: - Data is one of important asset of user or as well any organization that is moving on cloud. Data privacy issues are of great concern while moving data.
- e) Confidentiality: - a cloud provider should maintain the secrecy of communication between a provider and cloud user and different actions performed in various activities. This is measured by confidentiality parameter techniques which are used for achieving privacy of data message.
- f) Communication: - It measures strength against eavesdropping or leakage when data or messages passed in cloud computing environment. This parameter measures the provision provided by cloud service at time of data or message transmission.

- g) Isolation: - Problem of isolation of resources among multiple users arises from multitenant feature of cloud computing infrastructure.

III. TRUST MODEL ARCHITECTURE

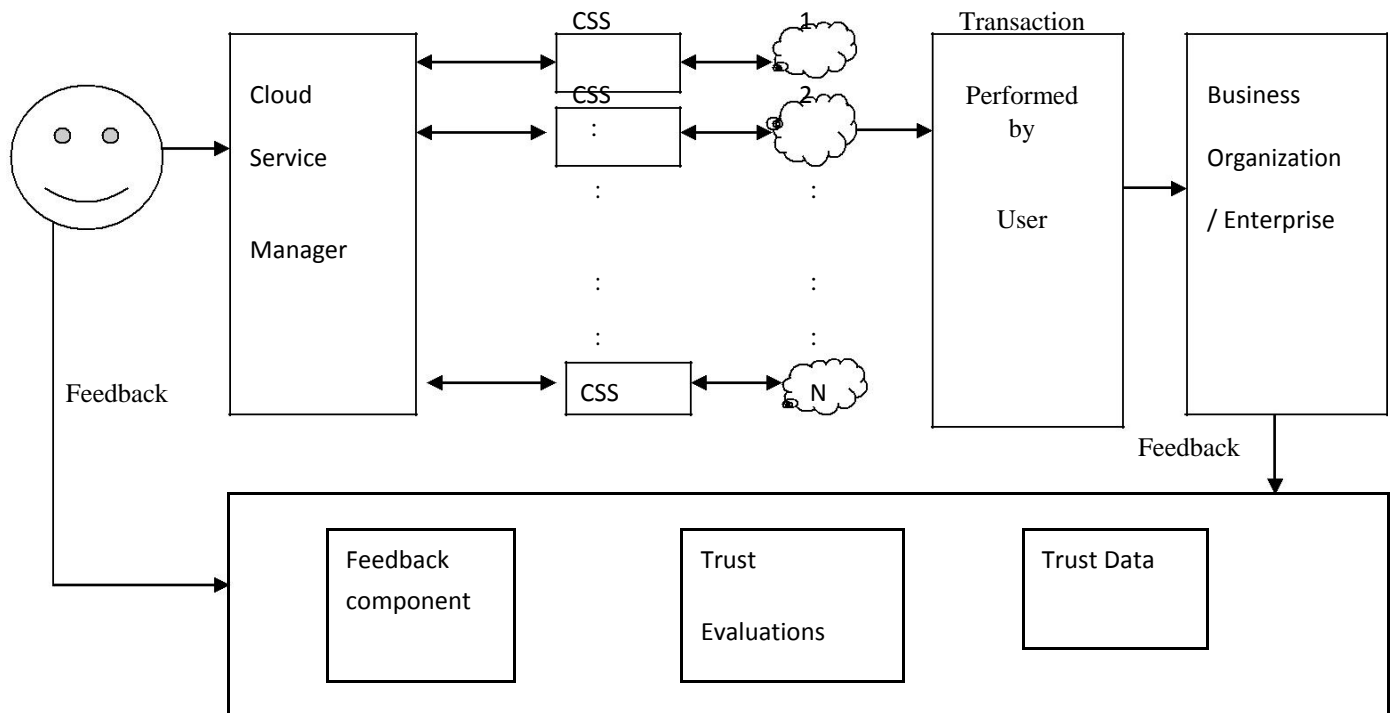
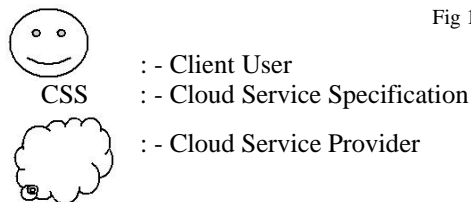


Fig 1: Trust Model Architecture.



Trust Model Architecture [2, 3] represents the role of trust between client and cloud service provider. Above diagram represents the transaction performed by the client such as online shopping etc. from the vendor who has opted for cloud computing. In above diagram client choose services from second service provider.

Components of Trust Model Architecture:-

1. Cloud Service Manager: - This component is responsible for maintaining information about a specific cloud service such as type of service, service provider and number of users registered for that service and other accounting information. It also maintain trust value for a particular cloud service that gives its security strength[21] .It is necessary for a cloud service to get registered with cloud service manager before its use. At this point static trust calculation is done. Previously, Trust value is calculated using static parameters with passage of time dynamic trust value can be evaluated. It provides opportunity for a user to select a particular cloud service.

2. Cloud Service Specification: - It maintains the specifications provided by particular cloud service provider. This detailed information is recorded by cloud service manager.
3. Cloud Service Provider: - Cloud Service Provider provides platform for various cloud computing services such as Infrastructure as a Service, Software as a Service, Platform as a Service, storage etc. It provides services through any cloud network it may be public, private or hybrid and charges bill accordingly known as utility bill.
4. Transaction: - It is element for which cloud is meant for such as transaction performed by user to purchase an electronic good from the vendor who has opted for cloud computing.
5. Trust Model has three components
 - a) Feedback Collector: - It collects the feedback provided by client as well as business organization. This feedback may contain challenge faced by both organization and client regarding trust. It helps further to improve the trust model.
 - b) Trust Evaluation:-Trust is evaluated based on various parameters such as authorization, authentication, and confidential data etc.
 - c) Trust data: - It contains the database of trust value associated with particular service provider.

IV. TRUST MODELS

This section reviews various trust models proposed for web services. Different parameters on which trust model for single web are listed below [4,8,10,11]

- (i) Parameter 1(P1) QoS Metrics: - It should cover multiple QoS metrics such as response time, availability, throughput, etc. to enable users well distinguish among functionally similar services.
- (ii) Parameter 2(P2) User Preferences: - It can vary from individual to individual. One may be interested in quality metric another in some other means.
- (iii) Parameter 3(P3) Feedback from Users: - Feedback from users which may be subjective or objective perspectives.
- (iv) Parameter 4(P4) Credibility of raters to avoid collusion and deception.
- (v) Parameter 5(P5) Bootstrapping mechanism is required to compute initial trust value for newcomer web services.
- (vi) Parameters 6(P6) A trust dynamism criterion is necessary to evaluate because performance of web services is subject to change over time.
- (vii) Parameter 7(P7) Dependency between credibility of majority of ratings and recommendations given to certain service should be avoided. The following table shows dependence of the model on different parameters for calculating trust value such as feedback model depends on parameter P1,P2 ,P3,P4,P5,P6, P7 etc.

| Model | P1 | P2 | P3 | P4 | P5 | P6 | P7 |
|----------------------------|----|----|----|----|----|----|----|
| a) Feedback based model | Y | Y | Y | | | Y | Y |
| b) Statistics based model | Y | Y | | Y | | Y | |
| c) Fuzzy logic based model | Y | Y | Y | Y | | | Y |
| d) Data mining model | Y | Y | | | | | Y |

Table 1: Comparative Summary for dependence of parameters of various models.

- (i) Feedback based Model: - Feedback based models build trust value from the reviews collected by either a provider or consumer about particular web service. Provider analyzes the service in the service registry and consumer generated information on basis of online reviews provided by users during past interactions. The ratings are based on QoS metrics such as latency and throughput but may also involve user preferences , feedback from users , trust dynamics and dependency between credibility of ratings and recommendations given to certain service. It has advantages of considering the opinions of users, which tend to be more rationale and meaningful metric for building any service. It also suffers from few limitations such as no bootstrapping mechanisms for computing trust for initial values. Second, Quality and credibility of ratings are major problem with feedback model because consumers view services from different perspectives according to their need for service. Some malicious raters may also give false rating.
- (ii) Statistics based model: - Statistics based model used to describe the relationship among set of variables with the help of mathematical equations. This led to improvement over feedback models which only depend upon the reviews provided by provider and consumer that may not be so accurate. It calculates trust value by combining multiple sources using statistical methods to combine them.
- (iii) Fuzzy logic based model: - It is reasoning approach that provides approximate rather than exact values. Fuzzy model helps us to analyze the semantic and rationale behind the feedback. On the basis of aggregating the feedback given by user on the same preferences together. Detect false rating feedback provided by malicious users for eg. User A always give false positive feedback instead of service perform badly in multiple invocations.
- (iv) Data mining model: - Data Mining approach used for extracting hidden patterns from huge datasets. However, for calculating trust a data mining based approach uses text mining to analyze the reviews provided by users in order. But in-depth methodology of text mining technique cannot be effectively performed. Moreover, trust dynamism and bootstrapping issues are ignored.

V. CONCLUSION

Cloud computing is an emerging technology but still it faces a lot of challenges and trust is one of them. In this paper, we have discussed trust model architecture and various parameters which are involved in calculating trust and used by trust models for single web service. Various trust models are discussed with their advantages and limitations. Trust models focus on reviews provided by consumers, providers, text mining technique is also used but models should also deal with bootstrapping mechanisms, penalize the dishonest behavior etc. In future research should focus on these issues.

REFERENCES

- [1] Md. T. Khorshed, A.B.M. S. Ali, S. A. Wasimi, "A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing", *Future Generation Computer Systems*, Vol.28, pp. 833-851, January 2012.
- [2] R. Shaikh, M. Sasikumar, "Trust Model for Measuring Security Strength of Cloud Computing Service", *Procedia Computer Science*, Vol. 45, pp.380 – 389, 2015.
- [3] S.K. Chong, J. Abawajy, M. Ahmad, I. R.A Hamid, "Enhancing Trust Management in Cloud Environment", *Procedia - Social and Behavioral Sciences*, Vol. 129, pp. 314 – 321, September 2015.
- [4] O.A. Wahab, J. Bentahar , H.Otrok , A. Mourad, "A survey on trust and reputation models for Web services: Single, composite, and communities", *Decision Support Systems*, Vol. 74, pp 121-134, April 2015.
- [5] A. J ,Sang, R. Ismail, C. Boyd, "A survey of trust and reputation systems for online service provision", *Decision Support Systems*, Vol. 43, no.2, pp.618– 644, 2015.
- [6] I.M.Abbadi, M. Alawneh, "A framework for establishing trust in the Cloud", *Computers and Electrical Engineering*, Vol. 38, pp. 1073–1087, July 2012.
- [7] S. N S. Kotikela, M. Gomathisankaran, "CTrust: A framework for Secure and Trustworthy application execution in Cloud computing", *International Conference on Cyber Security*, 2012.
- [8] H.T. Nguyen, W. Zhao, J. Yang, "A trust and reputation model based on Bayesian network for web services", *IEEE International Conference on Web Services*, IEEE 2010, pp. 251–258.
- [9] S. Phoomvuthisarn, "A survey study on reputation-based trust mechanisms in service oriented computing", *Journal of Information Science and Technology*, Vol. 2, no. 2, pp. 1–12, Dec 2011.
- [10] S.Vavilis, M. Petkovi, N. Zannone, "A reference model for reputation systems", *Decision Support Systems*, Vol. 61 pp.147–154, 2014.
- [11] Z. Malik, A. Bouguettaya, "Reputation bootstrapping for trust establishment among web services", *IEEE Internet Computing*, Vol .13, no. 1, pp.1089– 7801, 2009.
- [12] Y. Zhu, J.Wen , M. Qin, G. Zhou, "Web Service Selection Mechanism with QoS and Trust Management", *Journal of Information & Computational Science*, Vol.8,no. 12 pp.2327–2334, 2011
- [13] E. M.Mohamed, H. S Abdelkader, S. EI Etriby, "Enhanced Data Security Model for Cloud Computing", *8th International Conference on Informatics and Systems(INFOS)*, Cairo, pp.12-17, May 2012.
- [14] P. Gupta, M. K. Goyal, P. Kumar, A. Aggarwal, "Trust and Reliability Based Scheduling Algorithm for Cloud IaaS", *Proceedings of the Third International Conference on Trends in Information, Telecommunication and Computing*, *Lecture Notes in Electrical Engineering* Vol. 150, pp 603-607, 2013.
- [15] W. Fan , H. Perros, "A novel trust management framework for multi-cloud environment based on trust services providers", *Knowledge-Based Systems*, Vol. 70 ,pp. 392–406, July 2014.
- [16] K.J. Lin, H. Lu, T. Yu, C.E. Tai, "A reputation and trust management broker framework for web applications", in: *Proceedings of the IEEE International Conference on e-Technology, e-Commerce, and e-Service*, pp. 262– 269,2005.
- [17] K. Hwang and D. Li, "Trusted cloud computing with secure resources and data coloring," *Internet Computing*, IEEE, Vol. 14, no. 5, pp. 14 –22, sept.-oct. 2010.
- [18] M.K. Muchahari, S.K. Sinha, "A new trust management architecture for cloud computing environment", *International Symposium on Cloud and Services Computing (ISCOS)*, pp.136-140,17-18 Dec. 2012 .
- [19] S.Habib, S. Ries, M. Muhlhauser, "Towards a trust management system for cloud computing", *Trust, Security and Privacy in Computing and Communications (TrustCom)*,2011 IEEE 10th International Conference , pp. 933 –939, Nov. 2011
- [20] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud computing and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility," *Future Generation Computer Systems*, Vol. 25, no. 6, pp. 599 – 616, 2009.
- [21] R. Shaikh, M. Sasikumar, " Trust Framework for Calculating Security Strength of a Cloud Service", *2012 International Conference on Communication, Information and Computing Technology*, ICCICT 2012 ,19-20 October 2012,Mumbai, India.

Towards a mobile payment market: A Comparative Analysis of Host Card Emulation and Secure Element

Pardis Pourghomi
Department of Computer Science
American University of the Middle East
Kuwait

Pierre E. ABI-CHAR
Department of Computer Engineering
American University of the Middle East
Kuwait

Gheorghita Ghinea
Department of Computer Science
Brunel University London
Uxbridge, UK

Abstract— The considerable existing potential for mobile payments adoption shows that businesses are interested to increase the number of electronic transactions while consumers are attracted to convenient ways for fast and accessible banking. Nevertheless, the belief that the value of the Near Field Communication technology has not yet been fully recognized - particularly in the consumer marketplace - persists. However, the introduction of Android 4.4 operating system namely 'KitKat' has pushed the Near Field Communication (NFC) market towards Android devices with the recently proposed Host Card Emulation (HCE) technology. Moreover, there are various debates about the ways in which mobile payment processes should be managed. Currently, the most recognized and accepted methods for managing the mobile payment processes are the traditional Secure Element (SE) approach and the Host Card Emulation which has lately become a crucial topic for key industry players. This paper describes the aspects of moving forward with mobile wallets. More specifically, a broad discussion is developed to consider the pros and cons of both approaches. Correspondingly, a detailed analysis is carried out centred on the security and adoption issues that these approaches may raise.

Keywords- *Near Field Communication; Secure Element; Host Card Emulation; Mobile transaction.*

I. INTRODUCTION

Despite leading the mobile payments market in the world, the U.S. is nowhere near Europe when it comes to alternative card payment methods to the old swipe and sign system. Chip card is the technology currently being used for credit/debit card payments in most European and other developed countries around the globe. In order to carry out a transaction, consumers should insert their credit/debit card into the Point Of Sale (POS) device and enter their Personal Identification Number (PIN). With embedded chip payment method, consumer's signature is not required as a correct PIN would prove to the POS terminal that the payer is the genuine cardholder. This payment method provides an easy and secure customer authentication as well as payment authorisation for both payers and payees. The consumer PIN is stored in the embedded chip which tells the POS terminal whether a correct number has been entered in order to enable the POS terminal

to first authenticate the payer and then to authorise the transaction. The nature of this method of making payments is similar in all countries that run this service but branding might be different. For instance, this technology has been introduced as Chip and PIN in the United Kingdom and Ireland but in other European countries it is referred to as Europay, MasterCard and Visa (EMV). In addition to MasterCard and Visa, EMV is also supported by American Express and Diners Club [4].

Since the initial contactless technology and the advanced mobile payments methods are developed based on the EMV system, the wide spread of EMV has caused major delays in the adoption of Near Field Communication (NFC) mobile payments. This is due to the fact that each involved party claims to have more control privileges over the customer credentials than the other. This is in addition to the need for mobile network operators to play an enormous part in cooperating with financial institutions in order to manage the NFC mobile transaction processes [17][21]. On the other hand, since the card payment industry in the U.S. is still based on magnetic stripe and customer signature scheme, service providers bypass issues such as dealing with the ubiquitous of EMV systems and therefore the adoption of mobile payments have become considerably easier and faster than in Europe. Having said that, according to the Wall Street Journal [29] the EMV system has recently found its way to the U.S. and as a consequence it has become compulsory for all vendors to switch their existing swipe and sign systems to Chip and PIN by October 2015. This is done by MasterCard and Visa through the shifting the fraud liability from themselves to merchants who still use swipe and sign method when customers possess chip cards. Furthermore, The Wall Street Journal explains that "The U.S. is the last major market to still use the old-fashioned swipe-and-sign system, and it's a big reason why almost half the world's credit card fraud happens in America, despite the country being home to about a quarter of all credit card transactions".

Our goal is to provide a concept for the mobile payment ecosystem that is technically feasible, is accepted by all parties involved and thus provides a business case for each player in this ecosystem. The purpose of this paper is thus to provide a

balanced analysis of both HCE and SE approaches for financial institutions and mobile network operators looking at NFC payments. In so doing, we focus on the following research question:

“Which of the Host Card Emulation and Secure Elements approaches is better suited for an NFC-based mobile payments ecosystem?”

To this end, we employ a comparative methodology to study the relative strengths and weaknesses of both approaches from a four-pronged perspective: security, benefit to the NFC ecosystem, maturity, and usability. Comparison is implicit in all sciences and the use of the comparative method as an alternative to more traditional methods of scientific enquiry (e.g. experimental, case study) is by now long been established [11].

Accordingly, the rest of this paper is structured as follows. Section 2 provides an overview of the concept of Host Card Emulation. Section 3 discusses the potential security issues in mobile payments. Section 4 delivers a comprehensive evaluation of both HCE and SIM SE approaches to establish a foundation for further analysis. Section 5 provides our analysis based on multiple criteria. Section 6 highlights a number of research challenges which should be taken into consideration by stakeholders. Finally, section 7 presents conclusions and suggestions for future work.

II. NFC-BASED MOBILE PAYMENTS

Nowadays, mobile payments are using a number of different technologies to perform their transactions. Proximity payments rely on a contactless interface to chip-enabled payment technology, such as NFC-enabled mobile phones, HCE, SE, contactless stickers, etc. The mobile payments technologies including Google Wallet, Apple Pay, BellID, Simply Tapp, and Visa PayWave have many characteristics that are very similar, but they also differ in significant ways when it comes to implementation, security and user experience. In this section, we will provide a survey over few similarities and differences between these techniques.

In 2011, Google developers released Google Wallet v1.0, a mobile payment system that allowed its users to store debit cards, credit cards, on their mobile phone. Google Wallet v1.0 used to use the device-based Secure Element (SE) for card emulation and for making secure payments fast and convenient. According to google, this approach didn't work out well as most of the major network operators (Verizon, AT&T and T-Mobile) decided to support their own brand of wallet called Softcard and blocked access to the Secure Element for any other wallet providers. As an alternative, Google wallet v3.0 was developed and does not use a device-based Secure Element. It uses a technology called Host-based card emulation (HCE) instead, where card-emulation and the Secure Element are separated into different areas. In general,

this approach is considered less secure compared to the embedded SE approach. However, there are some areas (like Lost & Stolen use-case) where it is more secure.

Another approach was presented by Apple by combining both secure element and cloud-based technology into one framework. According to Apple, Apple Pay uses the secure element to store tokens and the payment client and adds biometrics with Touch ID for multifactor authentication. It allows Apple to use the power of local and backend data for risk management while removing all doubt about the security of a token or credential. Apple Pay does not store the real card data inside the SE. This is in direct contrast to Google Wallet 1.0 and Softcard. Moreover, Apple owns and controls the Secure Element embedded inside the device thereby avoiding unnecessary challenges from the MNOs.

Visa PayWave is an enhanced contactless payment application for cloud-based payments by introducing a new feature in the Android mobile operating system employing HCE, which allows any NFC application on an Android device to emulate a smart card, letting users wave-to-pay with their smartphones, while permitting financial institutions to host payment accounts in a secure, virtual cloud.

BellID technology enables mobile NFC transactions to be made by storing and accessing credentials in a remote environment rather than on the mobile devices. This gives application issuers independence and direct control to manage their credentials without any third party involvement. The software provides the functionality to complete transactions through a remote secure element (SE) using a common contactless terminal and acceptance infrastructure. When the consumer makes a transaction, their credentials are accessed from the remote cloud SE and passed through the mobile device to the point-of-sale terminal. The technology is compatible with and without HCE. According to BellID, their own application provides several features: offline transaction allowing consumers to make transactions even when no connection to the server is available; transaction optimization allowing transaction times to be minimized by the number of network round-trips for any application [2].

SimplyTapp is a card issuer solution provider for cloud based payments which provides mobile commerce services combining Android HCE with a cloud platform that makes commerce easy and secure. With the virtual secure element and host card emulation, SimplyTapp enables organizations to offer the convenience and flexibility of mobile NFC transactions without negotiating with device manufacturers and mobile network operators [15][24]. This solution, as well as the more general approach of performing NFC-based mobile payments using HCE technology, shall be explored in more detail below.

A. Host Card Emulation

Being able to perform NFC transactions through Host Card Emulation technology, Android 4.4 or KitKat has been

regarded as a highly significant platform. KitKat provides an environment for any NFC-enabled device such as phone and tablet to initiate and perform transactions with payment apps through the HCE technology that enables the device to emulate as a contactless smart card. HCE allows users to choose an app of their choice to make payments, which further helps them to avoid facing the potential issues with provisioning and personalisation [26] of the SE approach before and after the card issuance. The HCE features allow payment apps to effectively become a reader for NFC transactions through using a new Reader Mode. With HCE, the NFC-enabled device potentially acts as a contactless card that can be utilised for displaying IDs, making payments, showing tickets and vouchers. Most importantly, HCE stores customers' credit card details in a cloud environment which in this case enables the payment app to become a mobile wallet app. The proponents of HCE consider this feature as the most prominent part of this technology where a non ISIS [7] or Google Wallet [13] customer is similarly able to use his/her device for NFC purchases.

SimplyTapp was the first company which began rising the idea of managing remotely-operated smart cards in 2011. Before creating the term 'HCE', physical space was the only environment that card emulation could be achieved. More specifically, this means only a single party is able to replicate a card in a multi-purpose SE that is typically embedded in the Subscriber Identity Module (SIM), stored within the NFC phone. In August 2012, SimplyTapp publicly implemented the concept of HCE through launching Tapp NFC for CyanogenMod Android operating system [15][24] and up until the most recent implementation of HCE in KitKat, the service was only being offered by SimplyTapp. Figure 1 demonstrates an overview of this approach. The majority of NFC mobile payments services are however carried out based on the SE approach. This means that the card details are stored in a tamper-resistant chip on the phone that depending on the deployed ecosystem architecture is only accessible by certain apps. SE is also used by the mobile payment system to perform certain authentication tasks. Due to the lack of standardisation, there are several ecosystem architectures proposed for managing those apps.

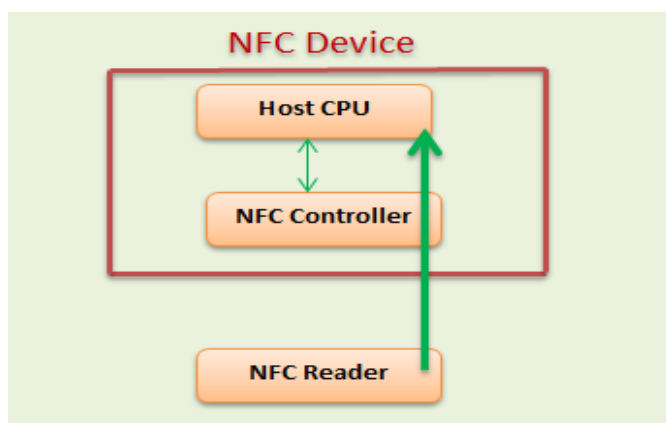


Fig. 1: HCE approach

One of the prospective models is when a network carrier is in charge of managing apps stored on the SE. For instance, Google Wallet only runs on a few carriers in the U.S., those are the only carriers that allow Google's app to access the SE whilst other network carriers block it on their phone [27]. There are some other U.S carriers such as AT&T, T-Mobile and Verizon that do not support Google wallet but rather have a rival system named SoftCard (formerly ISIS mobile wallet) [10]. The significant role of network carriers in the SE approach is bypassed through using the HCE technology, and the mobile payment providers such as Visa, MasterCard, Google, etc. can run their services on every NFC-enabled smartphone [16]. HCE also facilitates the opportunity for banks to offer their own NFC mobile payment services through their one-tap apps.

III. POTENTIAL SECURITY CONCERNS IN MOBILE PAYMENTS

Security in EMV transactions is provided by the embedded chip in the card. The chip as a tamper-resistant component ensures highest level of security through cryptographic checks whilst storing the transaction credentials as well as during the communication with the POS reader. Additionally, those cryptographic checks validate data integrity and client verification through checking the PIN [12][18]. Similarly in mobile payments, the *SE as part of SIM* approach provides the same level of security since transaction credentials are stored in a tamper-resistant chip that cannot simply be attacked. Thus, there is no Android application involved in the transaction process which could possibly introduce additional security concerns. However, the Android application might have a minor role in the entire NFC mobile payment process that could be the communication with the SE to receive information about the status of the transaction in order to notify the user of the app. This has nothing to do with processing the payment credentials as SE is the only responsible component in the phone to store, handle and process the transaction, and authentication data. In other words the SE acts like the smart chips in EMV cards.

However, in the HCE approach where the SE does not exist, an Android app plays a prominent role for providing the missing functionality of the SE that is receiving the NFC data in the transaction process. Nevertheless, the security properties of a tamper-resistant hardware chip is said to be greater than an Android app [5][8][9]. The proponents of the HCE approach argue that the security of NFC data is guaranteed since the controller directly transmits the credentials to the Android app and thus it would make it impossible for attackers to spoof a transaction app, containing data from a different source. Figure 2 illustrates the SE approach where there is no Android app involved in both data storage and transaction process.

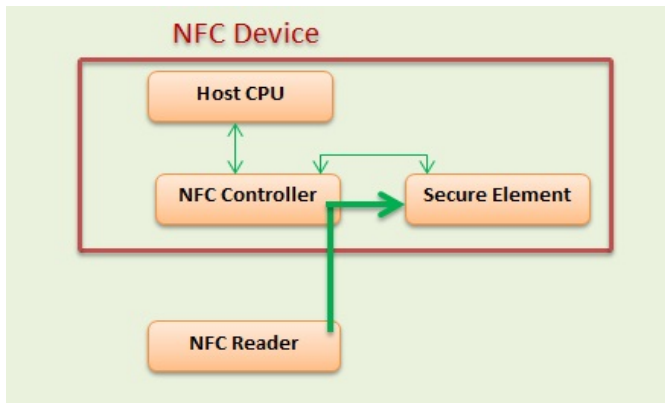


Fig. 2: SE approach

We believe the security of the Android app is provided through the Android application sandbox where transaction data and code execution is isolated from other apps on the device, however, we trust that there are certainly more security risks on a software application as opposed to a tamper-resistant hardware chip. To enhance the above addressed security risks and in order to improve the flexibility of NFC payments ecosystem, we proposed the NFC Cloud Wallet in 2012 [19][20] and suggested a virtual SE approach (SE in cloud) which had been then tracked and further developed by Visa, MasterCard and Bell ID. The idea of this approach was to store a virtual SE in the service provider's cloud and use any form of tamper-resistant chip in the handset to temporarily provide data storage in addition to handling the handset's authentication data. There are however different approaches towards implementing this model. We suggested having a GSM network to establish the link between the cloud and the NFC phone where the MNO is the cloud's owner while it fully manages the payment process. This is a so-called MNO-centred approach, where the financial institution is presumed to be part of the MNO. There are further alternatives, presently implemented cloud-based approaches which utilise the Internet channel for effective communications between the cloud and the NFC phone [30]. Yet, those approaches lack a single entity at the cloud end to lead the management of payment process rather than only dealing with app storage, which then raise its own obstacles. One of the difficulties of this approach is the need for Internet connection during transaction initiation, authorisation and execution. This would need the cloud provider to establish a connection with the merchant's POS terminal in less than half a second, which seem impractical. Consequently, companies tend to use payment tokens. Smart phones are where these tokens are locally stored with the aim of authorizing payment apps for carrying out transactions only for a limited number of days, for example. New tokens are fetched when a token reaches its limited number of usage or upon expiration of old tokens. As part of the regular sync process, tokens are managed when users are online or in the instance of offline users tokens are managed in the background. In this fashion, the security of the entire payment process is improved as the transaction information is useful only for a limited amount of time so that

the hacker is unable to perform an attack as long as s/he requires.

IV. RECOGNIZING THE RIGHT APPROACH

Early transaction models for mobile payments were primarily based on the idea of storing security information (e.g. transaction keys) from a debit card to a similar type of tamper-resistant card such as SE in the NFC device. The SE approach enables the NFC phone to emulate the card through securely storing MNO and Bank's keys in the MNO's Universal Integrated Circuit Card (UICC) [1]. Traditionally, the SIM had been managed by the MNO and the handset manufacturer was in charge of the embedded chip stored within the SIM. From the financial institution's point of view, the undefined integration of the SIM and the chip raised commercial issues which have been influential in delaying the acceleration of NFC payments. Despite barriers, the SE approach was established by the proponents since choosing to accept the security risks was not previously possible and contactless readers were only able to communicate with the SE and not the handset application. HCE enables the NFC device to receive and store contactless data on its own memory rather than a secure tamper resistant chip such as SE. In order to enhance the security of contactless data, with HCE, only a limited group of transaction data that can be used (e.g. for single transaction or merchant-specific data) is stored into the NFC device. That data is a so-called token which can be updated and deleted as the issuers require.

In the world of NFC payments, the idea of relying on the SE has not only been an issue in the U.S. marketplace but it has been and still is considered as a distinct problem in every market. Principally, having to rely on the SE requires network carriers' participation in the whole transaction process. This would not be the ideal solution for stakeholders since the most practical form of this approach is processed through the SE as part of the SIM schema that allows the network operator to get extensive access to the transaction information. Storing contactless data in the actual mobile device may be an alternative approach whereby the need for cooperating with the network operator is replaced by the handset manufacturer or the chipset maker.

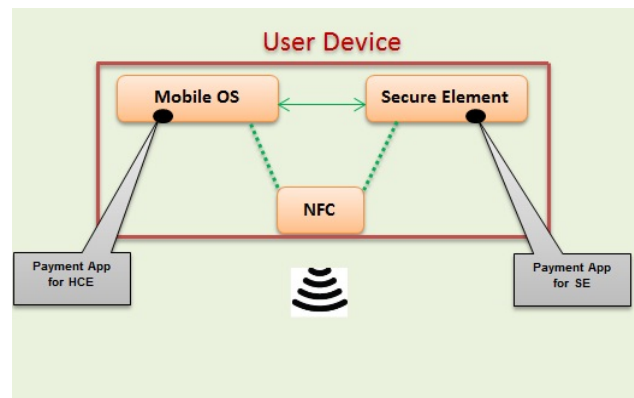


Fig. 3: SE and HCE in the handset

NFC standards enable this technology to run across multiple operating systems (e.g. Windows phone) while it is compatible with wide range of chipset card manufacturers. Hence, the enablers of HCE approach claim that SE in form of software is the best solution for bypassing the reliance on mobile network operators. Android 4.4 is an exact example of SE as software which has been released by Google. It supports HCE technology which allows communication between any NFC enabled device and contactless payment terminal through emulating a physical contactless card.

A. The SIM SE NFC ecosystem

Figure 4 illustrates the ecosystem we initially proposed, developed and described in [13] as a basis for a detailed exploration of the traditional SE approach. We thus designed a number of secure transaction authentication protocols in [19][22][31] to support our “NFC cloud wallet model” SIM SE based NFC payments in which the MNO plays the main role for managing the transaction process and is integrated with the financial institution. The payment app contains transaction credentials, performs contactless Europay, MasterCard and Visa (EMV) transaction with the POS and is stored in the SE that is part of the SIM (e.g. UICC). This payment app appears as a contactless payment card to the POS terminal. Additionally, the user interface is provided through an app that is stored in the NFC device but outside of the SE which enables the user to communicate with the payment app in order to carry out several activities such as check balance, enter passcode, etc. The issuing bank (payment app owner) should provision the app into the SE before allowing any transactions to occur. This would not be possible without the cooperation of the MNO in the SIM SE approach who will grant the issuer access to the SE. Once completing the provision stage, the issuer can access and manage the credential stored within the SE and in the payment app. The ‘A’ indicates the first approach in which the MNO and merchant communicate through the NFC phone. From the protocol design point of view, there is not a shared secret key between the MNO and the merchant. The second approach shown as ‘B’, facilitates a method for both the MNO and merchant to communicate directly and without the

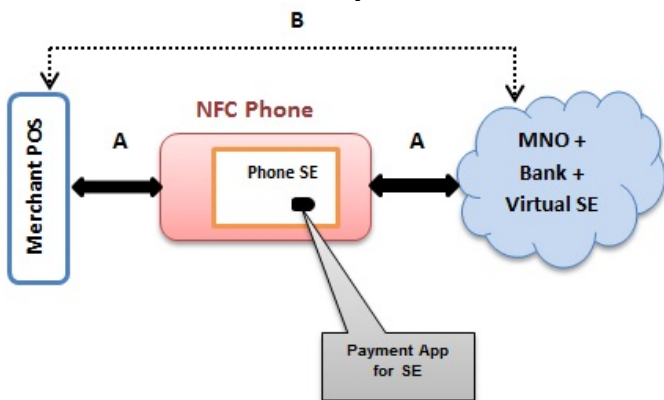


Fig. 4: The SE ecosystem

involvement of the NFC phone that is the actual consumer. The prerequisite for this approach is a shared secret key which should be sent and received by both MNO and merchant. As mentioned above, our proposed secure transaction authentication protocols which cover both approaches comprise a comprehensive security analysis for validating the reliability of such scenarios as well as their protocols to prove the resiliency of the protocols against malicious attacks.

B. The HCE NFC ecosystem

Figure 5 illustrates the ecosystem which was first developed by SimplyTapp [15][24] to support HCE-based NFC payments. This approach provides a number of strategic differences in comparison with the conventional SE-based NFC payments, which are described in what follows.

In the HCE ecosystem approach, the use of SE is avoided and replaced by a payment application which contains transaction credentials and resides in the NFC device operating system. In this case, the mobile app appears as the payment card to the POS terminal and similar to other apps can be downloaded from an app store directly to the NFC device. The need for the involvement of a MNO is avoided since HCE offers more flexibility by not using an SE, however, the security and usability of the payment app that resides in the handset’s operating system remains as the main concern. Naturally, a mobile based payment app does not offer the same level of security as a hardware SE which consequently raise the need for additional methods to be added to its operation.

For instance, HCE providers may decide to provision single use or limited use transaction credentials to the payment app or the mobile device itself to minimise the impact of a malicious activity. No long term storage of sensitive information, such as Personal Access Number (PAN), Personal Identification Number (PIN) and keys, is appropriate within the mobile handset operating system. The need to provision limited use payment credentials increases the possibility of frequent usability issues as well. For instance, without the presence of Internet connectivity the mobile phone won’t be able to receive a new payment token from the issuer bank.

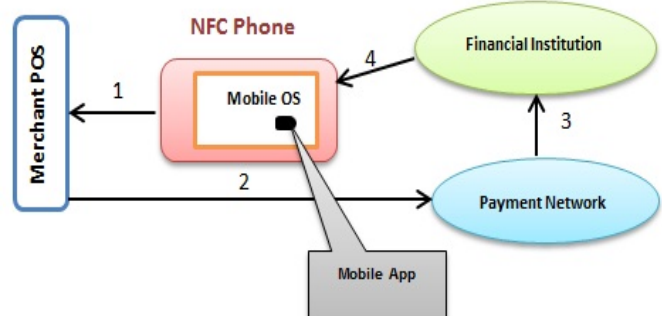


Fig. 5: The HCE ecosystem

V. ANALYSING BOTH APPROACHES

Since the arrival of Google's KitKat that supports HCE solution, there have been various debates and views about the scheme of processing NFC mobile wallets. The main focus of those debates is yet around selecting either the HCE contactless or SE based transactions models.

The level of flexibility that HCE brings to stakeholders is undoubtedly more than what the traditional SE approach (where MNO and bank are disintegrated) has to offer. In order to process the transaction, the service provider does not need to rely on the SE or the Trusted Service Manager (TSM) which itself arguably makes the route to market smoother and cheaper. However the superiority of the SE approach, either SIM SE or embedded SE, is proven to MNOs and handset manufacturers. In fact, the operation mechanism of the two approaches is very different in terms of involved entities, security, usability, etc. nonetheless, we believe that looking at these two approaches as competitors won't assist the existing concerns but rather these approaches should be incorporated as one payment processing mechanism.

The way in which HCE solution significantly reduces the payment process is through using tokenisation rather operating based on a tamper-resistant chip. However, tokenisation is considered as a newly introduced paradigm which has not yet satisfied its surrounding security risk issues. SIM secure element approach however recommends an MNO-centred model where the network operator is in charge of provisioning and managing the SE over-the-air. In contrast, embedded SE has always been an interest to handset manufacturers were they can take a bigger role since the SE is embedded in the handset and not in the SIM and as a result there is not much expected from MNOs. Generally, the SE solutions are generally recognized to be more secure, but perhaps less granular in control and they do need a more complex and binding ecosystem to deliver them. Having said that, the growth of the society is dependent on the technology and while HCE serves the NFC ecosystem, the technology remains with no standards and connected to the SE-based deployments.

A. Security and appropriate usage

In comparison to the Android application, HCE has no differences since it is not identified with any hardware or software for security services. It is well known by industry reports that Android is the most attacked of all mobile environments [3]. HCE is not part of the secure management used in the current mobile devices, neither it is deployed in the frameworks for SE. HCE does not enjoy the same ability, secure execution, remote management and the management of the mass roll-out and post-deployment of NFC applications in secured method with stable standards. We believe that HCE is best used in scenarios where the user's credentials are of low value and the user's privacy is not a requirement.

B. Benefit to NFC ecosystem











The purpose of HCE is to help the service providers and the developers to reach their aims by providing the market with new NFC services. It is expected by, SIM alliance, that new creativity in the field of the NFC ecosystem will be brought out to light. This will not cancel the traditional smart cards use but will help the players to develop the applications. This will come true by creating new NFC use cases and strengthening the NFC service of the present users. By the end, the purpose of this innovation is to increase the number of the applications as well as to make the NFC more accessible and familiar to end-users.

C. Certification schemes

SE security is empowered to face malicious operations. This advantage is remarkable for the SE as a standalone component. Knowing that the security is the major characteristic of SE, it cannot resist the most sophisticated attacks. As a result, there are no demonstrated instances of unauthorised access to, or duplication of the sensitive data stored in a SE. It is also a reliable 'second factor' in models of strong authentication where the first factor is 'something you know' and the second factor is 'something you have'. The support of the SE comes from the European Telecommunications Standard Institute (ETSI), the 3rd Generation Partnership Project (3GPP), and the Global Platform and the Java Card standards [6][25]. That explains the use of SE for the vast majority of services provider's requirements including banking, transport, access control, identity, etc.

To adopt NFC services for mobile it is a necessity to guarantee the security of service provider for end-user credentials [14][28]. The SE's certification scheme is greatly drawn.

Table 1: SE vs. HCE

| | HCE | SE |
|-----------------------|---|---|
| Business model |  |  |
| Security |  |  |
| Provisioning |  |  |
| Usability |  |  |
| Maturity |  |  |

Certifications can be accomplished after a deep analysis and tests conducted by a number of independent certification laboratories. As HCE is a new payment solution technology, it is very important to protect the payment schemes by identifying and developing countermeasures to possible threats to HCE, leading to the definition of a certification scheme that is sufficiently comprehensive to demonstrate the robustness of the solution to service providers and end-users. The fact that HCE is not a standalone piece of hardware but an integrated component in the Android OS (together, potentially, with other OSs in the future) is a serious challenge; it may be more difficult to identify the possible attack paths than it has been with the SE. Establishing the boundaries of stakeholder security responsibility is likely to be another big and related challenge.

The SIM SE or HCE choice has a commercial background; in practical terms, this means that the bank will need to choose between known costs with the lower delivery risk of SIM SE and the promise of lower costs, but with the chance of unexpected delivery costs and increased delivery risk of HCE - knowing that SE and HCE are not mutually exclusive, every bank should develop services to build competencies.

VI. RESEARCH CHALLENGES

It is important to highlight that a secure component in the device together with a secure provisioning and management solution, remains a necessary requirement for NFC mobile payments. This requirement serves the distribution and the management of valuable and sensitive credentials. It is also important that the secure component receives the applications in a "black-box" and enjoy robust software that cannot easily be removed, destroyed or hacked.

We believe that it is necessary to have such a secure component and management system certified following extensive security testing by several recognised third-party laboratories, thus ensuring that the secure NFC ecosystem is audited using the latest generation of known attack path techniques. The only remaining issue for the mobile operator community is to enhance its cooperation with other involved participants of the ecosystem to simplify the provisioning process and globally accelerate the deployments of SIM-based NFC.

In addition to the above, we also suggest a number of key points for banks to consider as they plan mobile NFC payments:

- Study the local environment to determine the best approach
- Dress the target transactions: HCE might be less suitable to some transaction such as offline, high value than SIM SE.
- Design a flexible framework to overcome the intersection between SIM SE and HCE concerning the required systems and capabilities.

- Co-operate with the e-business industry such as banks sector to provide the required level of standardisation for HCE.
- Understand the online risk management system; handling risk management for online payments systems has become a part of the electronic payment landscape. However, the security will not, at least in the short term, achieve the recognized levels of SIM SE-based security and may require providing a novel framework to handle these risks.

A key issue with storing the SE "in the cloud" is one of authentication. How does the issuing bank know that the "SE in the cloud" is being accessed by a legitimate device and user? Without the use of the SE, the issuing bank is likely to need to employ a range of measures (such as device fingerprinting and risk based authentication) to maximize their chances of detecting unauthorized access. Mobile security capabilities will be required as well to detect devices that have been rooted or compromised in some other way. These capabilities will need to support a range of devices and changing OS versions.

VII. CONCLUDING REMARKS

This paper provided a comparison of HCE with the early SIM SE approach for managing mobile payments ecosystem. Consequently, it found that although HCE simplifies the complexities of managing mobile payments ecosystem (e.g. application provisioning), it still raises a number of security issues which should be considered the priority from the technology viewpoint. At present, in order to deploy the HCE approach, financial institutions should cooperate with service providers and payment network operators to get hold of certification waivers and ultimately deliver the service. The cooperation would need to be in place until the rules for HCE are completely designed and developed. On the contrary, the SIM SE approach for managing mobile payment processes provides a mature but complex method that enhances the security of sensitive information while making the collaboration of involved parties more complex in the payment ecosystem. As a result of this complexity and due to the critical role that mobile network operators play in provisioning mobile payment services, network carriers are actively working to simplify this process.

Returning to the research question posed at the outset of this article, we can say that, despite the fact that HCE approach provides simplicity and flexibility, to ensure data security, we nonetheless suggest that financial institutions should carefully review their needs and follow the SIM SE approach in cooperating with service providers and network operators in order to deliver a reliable and effective service.

The work done up until now in the field of mobile payments indications that utilising SE for data storage as well as in carrying out NFC mobile transactions results in an interoperable environment which can accelerate the mass

deployment of NFC applications. Since security has always been the main concern in mobile payments, the SE as a tamper-resistant component does indeed offer the most stringent level of security to date while it is supported by mature certification schemes. As part of future work, the usability issues with HCE would need to receive a particular attention as the use of this technology matures.

REFERENCES

- [1] Alimi, V. and Pasquet, M. 'Post-distribution provisioning and personalization of a payment application on a UICC-based Secure Element'. In *IEEE International Conference on Availability, Reliability and Security (ARES'09)*, pp. 701-705, 2009.
- [2] BellID, 'Host Card Emulation', 2015. Available at: <http://www.bellid.com/mobile/host-card-emulation/>
- [3] Cisco, '2013 Cisco Annual Security Report', 2013. Available at http://www.cisco.com/web/offer/gist_ty2_asset/Cisco_2013_ASR.pdf
- [4] EMV Migration Forum, 'Implementing EMV at the ATM: Requirements and Recommendations for the U.S. ATM Community', 2014. Available at: <http://www.emvconnection.com/wpcontent/uploads/2014/08/EMV-Implementing-EMV-At-The-ATM-20140813-FINAL.pdf>.
- [5] GlobalPlatform, 'Secure Element Access Control v1.0.20', 2014. Available at: http://www.globalplatform.org/specifications/review/GPD_SE_Access_Control_v1.0.20.pdf
- [6] GlobalPlatform, 'Card specifications', 2014. Available at: <http://www.globalplatform.org/specificationcard.asp>
- [7] GSMA, 'Case Study: ISIS Mobile Wallet', 2014. Available at: <http://www.gsma.com/digitalcommerce/wp-content/uploads/2013/10/Mobile-Wallet-Case-Study-ISIS-digital-version1.pdf>
- [8] Handford, R. 'Report: HCE vs SIM secure element is not 'either or' choice', Mobileworldlive, 2014. Available at: <http://www.mobileworldlive.com/report-hce-vs-sim-secure-element-either-choice>
- [9] Janssen, T. and Zandstra, M. 'HCE security implications – Analyzing security aspects of HCE', 2014. Available at: <http://www.ul-ts.com/downloads/whitepapers/finish/6-whitepapers/289-hce-securityimplications>
- [10] Kharif, O. 'AT&T-Verizon-T Mobile Sets \$100 Million for Google Fight: Tech', Bloomberg News, 2011. Available at: <http://www.bloomberg.com/news/articles/2011-08-29/at-t-verizon-t-mobile-sets-100-million-for-google-fight-tech>.
- [11] Lijphart, A. 'Comparative politics and the comparative method. *American political science review* 65(3):682-693, 1971.
- [12] Madlmayr, G., Langer, J., Kantner, C., and Scharinger, J. 'NFC devices: Security and privacy'. In *IEEE International Conference on Availability, Reliability and Security (ARES 08)*, pp. 642-647, 2008.
- [13] Moss, A. 'Google Wallet Implementation and Looking Beyond', VeriFone, 2012. Available at: <http://global.verifone.com/company/resources/>
- [14] Munch-Ellingsen, A., Andersen, A., and Akselsen, S. 'Manage your own security domain on your smartphone'. In *First IEEE Conference on Mobile and Secure Services (MOBISECSERV)*, pp. 1-7, 2015.
- [15] NFC World, 'SimplyTapp', 2012. Available at: <http://www.nfcworld.com/simplytapp/>
- [16] NFC World, 'MasterCard backs host card emulation for NFC payments', 2014. Available at: <http://www.nfcworld.com/2014/02/19/327936/mastercard-backs-host-card-emulation-nfc-payments/>
- [17] Ondrus, J., and Pigneur, Y. 'An assessment of NFC for future mobile payment systems'. In *IEEE International Conference on the Management of Mobile Business (ICMB 2007)*, pp. 43-43, 2007.
- [18] Pasquet, M., Reynaud, J., and Rosenberger, C. 'Secure payment with NFC mobile phone in the SmartTouch project'. In *IEEE International Symposium on Collaborative Technologies and Systems (CTS 2008)*, pp. 121-126, 2008.
- [19] Pourghomi, P., and Ghinea, G., "Managing NFC payment applications through cloud computing." *Internet Technology and Secured Transactions, 2012 International Conference for*. IEEE, 2012.
- [20] Grønli TM, Pourghomi P, Ghinea G. Towards NFC payments using a lightweight architecture for the Web of Things. *Computing*. 2014;1-5.
- [21] Saeed MQ, Pourghomi P, Walter C, Ghinea G. Mobile Transactions over NFC and GSM. In 8th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM). IARIA 2014 (pp. 118-125).
- [22] Pourghomi, Pardis, and Gheorghita Ghinea. "A proposed NFC payment application." *arXiv preprint arXiv:1312.2828* (2013).
- [23] Schamberger, R., Madlmayr, G., and Grechenig, T. 'Components for an interoperable nfc mobile payment ecosystem'. In *5th IEEE International Workshop on Near Field Communication (NFC)*, pp 1-5, 2013.
- [24] SimplyTapp, 'Mobile Payments as Easy as HCE', 2014. Available at <http://simplytapp.com>
- [25] Smart Card Alliance Mobile and NFC Council, 'Mobile/NFC Security Fundamentals Secure Elements 101', 2013. Available at: http://www.smartcardalliance.org/resources/webinars/Secure_Elements_101_FINAL3_032813.pdf
- [26] Smart Card Alliance Mobile and NFC Council, 'Host Card Emulation (HCE) 101', 2014. Available at: <http://www.smartcardalliance.org/wp-content/uploads/HCE-101-WP-FINAL-081114-clean.pdf>
- [27] Smith, J. 'Verizon is Blocking Google Wallet on Galaxy Nexus', 2011. Available at: <http://9to5google.com/2011/12/05/verizon-is-blocking-google-wallet-likely-because-of-isis-partnership/>
- [28] Tan, G. W. H., Ooi, K. B., Chong, S. C., and Hew, T. S. 'NFC mobile credit card: the next frontier of mobile payment?' *Telematics and Informatics*, 31(2), 292-307, 2014.
- [29] The Wall Street Journal, October 2015: The End of the Swipe-and-Sign Credit Card', 2014. Available at: <http://blogs.wsj.com/corporate-intelligence/2014/02/06/october-2015-the-end-of-the-swipe-and-sign-credit-card/>
- [30] Yang, J. H., and Lin, P. Y. A mobile payment mechanism with anonymity for cloud computing. *Journal of Systems and Software*, 2015.
- [31] Pourghomi P., Saeed MQ, Ghinea G., 'A Secure Cloud-Based Nfc Mobile Payment Protocol' *International Journal of Advanced Computer Science and Applications (IJACSA)*, Vol. 5, No. 10, 2014

AUTHORS PROFILE

First Author Dr. Pardis Pourghomi is currently an Assistant Professor of Computer Science at the American University of the Middle East (AUM) in Kuwait. Prior to joining AUM, he worked as an Information Security Consultant in one of the largest UK's mortgage lenders called Coventry Building Society. Before that, he worked as a part-time lecturer at Brunel University London while acting as a subject-matter expert in a number of cyber security projects. He received BSc (Hons), MSc and PhD, all in Information Security, from University of East London (2009), Royal Holloway, University of London (2010) and Brunel University London (2014) respectively. He is an active researcher/author with publications in a number of conferences, journals and white papers. His current research interests revolve around applied security and privacy architectures especially in the field of smart cards. His main academic research interests include design of secure authentication protocols, mobile payments, NFC, communication architectures, transport ticketing security. Growing areas of interest include maximizing the role of Cloud in IoT, and security management.

Second Author Dr. Pierre E. ABI-CHAR was born in Halat, Lebanon. In 1998, Pierre E. ABI-CHAR received a Master Diploma in Physics from the faculty of Sciences at Lebanese University, Beirut, Lebanon. He received an M.S. degree in Computer and Communications Engineering, CCE, from University of Balamand (UOB) in 2000. In 2004, he received a Diplome d'Etudes Approfondies, DEA, degree in Network Security from University Pierre & Marie Curie-France in collaboration with ENST-France, Lebanese University, and INRS-Telecommunication-Canada. He received his PhD diploma in Networks and Information Security, in March 2009, from "Ecole Doctorale Informatique, Télécommunication et Électronique (EDITE)" at Pierre and Marie Curie University (Paris-6) Paris-France, with the collaboration of Telecom SudParis, Paris-France. Mr. ABI-CHAR research interests span the fields of Applied Cryptography, Network and Computer Security, Information and Coding Theory, Wireless Security, Context-Aware Security Technologies, Access Control, Privacy, Trust Management and Risk Assessment, Authentication Mechanisms, Routing Protocols, Intrusion Detection and Prevention Systems, Trust Models.

Third Author Dr. Gheorghita (George) Ghinea is a Reader in Computing in the Department of Computer Science, at Brunel University London. Dr. Ghinea's research activities lie at the confluence of Computer Science, Media and Psychology. In particular, his work focuses on the building trustworthy end-to-end systems. Dr. Ghinea leads a team of 8 researchers in these areas and he has over 200 publications in his research field. He consults regularly for both public and private institutions in his areas of expertise.

Performance Analysis of Sybil Decline: Attack Detection and Removal Mechanism in Social Network

Deepti S. Sharma
Department of CSE,
LKCT, Indore, M.P, India

Dr. Sanjay Thakur
Department of CSE,
LKCT, Indore, M.P, India

Abstract- Peer to peer system involves communication between the two directly connected hosts. As it supports the open communication medium it suffers from various security threats from remote computing elements. To avoid these threats various systems and integrated solutions are proposed over the last few years. One of such well known problem is Sybil attacks in which user performs unauthorized or malicious activities by creating the fake identities over the webs. One way to keep such threat away from the system is to design a centralized trusted authority system. Among its protection approaches Sybil Defender, Sybil Limit, Sybil Guard and Sybil Shield are some well known tools. After analyzing these tools and their respective mechanism we have found that they somewhere lack the associated trust computation issues. For improving such issues we have suggested a novel approach named as Sybil Decline. This paper gives a performance evaluation of suggested approach on some performance monitoring parameters. Extensive analysis and number of experiments are performed here to prove the results authenticity and effectiveness of the suggested approach.

Keywords- Sybil, Security, Peer to peer system, trusted authority

I. INTRODUCTION

Sybil attacks are the most dangerous threat for social network due to its uncertain nature and lack of information verification mechanism. Here the communicating nodes and their identities varies frequently and continuously distort the normal operation of the system. In its least difficult structure, a single malicious node claims to have many identities with an intention of upsetting the regular behavioral operations of the distributed system by pretending as if it is multiple nodes. For an instance, numerous distributed system use voting mechanisms to create consensus among users, and a false node with numerous Sybil identities can easily be outvote as a benign users in such a system, so that it can decide their outcome and affect the operations of the distributed system. To protect the system against this attack, there have been several attempts that could be classified as centralized and decentralized solutions. To provide the solution for preemption of such attacks can be described in two broad categories: centralized and decentralized [1]. The authentication controlling power of verifying

identities is given to multiply users which were earlier only for the single user. This authentication is aided with the detection of false positive and negative attacks research works on Sybil defense techniques.

In the centralized approach proposed solution for this is to have a centralized authority that can be used to provide digital credentials (eg. cryptographic keys) and bind them with the provided identify of the participating nodes in the system. The decentralized solutions to address the Sybil attack, on the other hand, replace the centralized authority with decentralized mechanism to limit a malicious node by her enforceable credentials and publicly verifiable resources [2]. For example, nodes in distributed systems oftentimes have physical resources that are limited in nature, such as processing capabilities, memory, addresses, and geographical location, among others, and these can be verified by other nodes in the distributed system to establish the identity of that node. A late course to address the Sybil attack makes utilization the social networks and their means of trust. Generally, this course has some basic components and attributes with the decentralized solutions for its task, as not requiring a centralized control for operations, and with centralized solutions, in its dependency on resources that are difficult to imitate. For a large number of solutions regarding this task, social networks are used for bootstrapping the trust and operations of the distributed system and just the nodes that are existing in the network (or just the ones which are the outcome of the social network's or distributed system's natural growth) are permitted into the operations of the appropriated distributed system [3]. To make these kinds of identities embroil into the social graph, the antagonist needs to form a great number of edges between itself and the remaining of the network graph, as a result represent the honest node, which is associated with the high cost.

High defense against security vulnerabilities or attacks demands complete controls at different working stack of its present infrastructure. Further reducing it might be said that, there are two areas that you need to consider [4]:

- (i) Infrastructure security involves controls (such as firewalls, perimeter security, etc.) relating to the nuts and bolts of the network environment, that will help protect it from external attack. These attacks include infrastructure threats like viruses and worms, but also less lethal ones such as spam. The focus of infrastructure security should be on preventing malicious activities which causes drops in the actual working of the system.
- (ii) Service security is even at higher risk associated with it. It involves defending the key system, application and information from both inappropriate access and illegal use. Identity access management is typically the best approach to protecting these key assets and services that your IT organization provides.

For getting the deeper view of above security primitives first the requirements needs to be understood clearly.

II. BACKGROUND STUDY

The Sybil attacks are most difficult security threat over the internet because the malicious node here claims many different identities for affecting the regular system operations. Creation of such identities is controlled by the adversaries who control the decision making of providing the services after the verification of originality of the user. One of these approaches is trust based evaluations where each user is analyzed on some trust factors before serving the complete control of operations to the requester. So many other works decompose existing Sybil defense schemes and demonstrate that at their core, the various algorithms work by implicitly ranking nodes on the basis of their connectivity to the trusted nodes [5]. Nodes that have better connectivity to the trusted node possess high rank and are considered to be more trustworthy. Similarly some of the researchers had depicted the endeavors to recognize, characterize, analyze and understand Sybil account activity in the Renren on line social network. Sybil attack can be defended by two ways: Sybil Identification and Sybil Tolerance. This work will evaluate the performance of previously suggested approach named as Sybil Decline which improving the problems associated with traditional detection is suggested.

It prevents irregular and unauthorized communication that leverages the difficulty in establishing and maintains relationships in the social network. It is based on ranking mechanism based on their perceived likelihood of being fake (sybil). Sybil decline is computationally efficient and can scale the graph with large number of nodes, as demonstrated by developed prototype. The ranking mechanism is based on the trust calculation using trusted model. In this step each user has to reach a unique trust value which is more than a specific threshold which is defined by

allocator policies. It gives an insight to user activities which is defined after trust verification. The above mentioned trust of every user is analyzed by its historical data access. It is based on user categorization and access control policies. Thus, while measuring the trust value, the allocator from trust must assign different weights to the different factors on the basis of which the system composite or final trust is calculated. The weights assigned will be based on the trust evaluation policy of allocator. If the different allocator assigns different weights for same task then the final value of the trust count will be different. The second reason is possibly applicable only when the allocator is human being and is completely subjective in nature one person may be more trusted than another based on individuals belief. So it needs a system which concentrates on both of these concepts, but the last-mentioned notion is acutely difficult to frame.

During the last few years there are so many approaches suggested to overcome the Sybil issues with social network. In the paper [6] SybilGuard is given which is used for reducing the influences and effects of forged identities. It maintains the edge based relationship for authorizing the users and its connected peers. In the paper [7] a statistical model is presented with a learning primitive for collaborative network named as latent community (LC) using Bayesian inference model. Evaluation of work shows experimentally that LC-based Sybil detector competes well with algorithms for the Sybil detection from the network security literature. The paper [8] gives control criteria for early stage detection of Sybil using a trust relationship model through distributed protocol. It also gives mechanisms to prevent the malicious influence of misbehaving nodes that do not perform the computational work. The work also proves strong evidence that Sybil-Control can be practically deployed. Another tool named as SybilLimit [9] gives an extension to Sybil Guard in edge evaluations, balance conditions benchmarking and independent instance executions for getting a better control over the Sybil nodes.

SybilShield [11] is one another tool works for multi-community social network and reduces the false detection rates by using the agent nodes. These nodes perform validation of random routes denies the request of suspected node. Some of the researchers had worked on analysis of historical records of the users for getting the mapping of user's behavior as Sybil. One of such approach is presented in [12] which use KD-Tree along with SICT and SICTF algorithms to solve the community mining issues of Sybil removal. Privilege Attenuation [13], PrivacyJudge [14] and Trust in peer-to-peer systems [15] are some other presented work based on trust based evaluations with accurate and effective analysis. In the paper [16] a social network based anonymous identity verifies STOR is proposed for protected transmissions. Mainly it depends on fuzzy

model to determine trust relationships between friends based on qualitative and quantitative social attributes. STor users thus select routers by taking into account their trust in those routers. While several attacks and attack scenarios have been documented, few studies have measured the actual deployment of such attacks and none of the documented countermeasures have been tested for compatibility with an already deployed network [17]. The paper shows that the world-wide deployed KAD network suffers large number of disbelieving insertions around shared contents and enumerate them. It detects the attack after analyzing the sharing of the peers' ID found around an entry after a DHT lookup.

To date, there is no comprehensive evaluation of how effective social network-based Sybil defenses are when the underlying assumption is completely or partially false in Sybil tolerance or Identification schemes. Thus for Sybil defense schemes to work well, all non-Sybil nodes need to form a single community that is distinguishable from the group of Sybil nodes. So for improved and timely identification of such attacks fake user entry needs to be identified in accurate manner. Thus this work proposes a novel model for such detection.

III. PROPOSED SOLUTION

The primary objective of this work is to have better accuracy and higher detection rates which can be made feasible by modifying the traditional structure of detection. In a way to achieve its goal we have suggested SybilDecline [18] which predicts the separation criteria for false and genuine identity categorization. If any of the communicating node is operating with multiple identities are found as suspected, then it must be blocked with all operations. Also the approach provides successful detection in early stages of system which prevents the complete affections or service drops. The suggested mechanism is step wise process starts from data generation and whose successful execution with all its monitoring parameters will lead towards the identification and removal of attacker. The data generation module will captures the activities of all the nodes connected and communicating with each other. The system analyzes the behavior of each node by their associated meta-data and their historical events. After analysis they separate the nodes into classes: Directly trusted and Suspected as Sybil. The model is a quantitative approach whose evaluation is completely numerical. Here the complete trust is calculated by assigning the initial weight values to different user activities which they perform towards their assigned services. The approach also made a check for trust deciding threshold conditions which limits the boundaries. It only allows the user having higher trust values than the decided threshold and the rest users are interrupted.

The prime concern about such user and community management is for large networks which stores tera bytes of data and which has to be processed frequently to get the intelligent business decisions. Community networks spreads the information very fast so as to identify the bogus or fake profiles is very complicated task. The behavior of the user can be used to further identify the accurate decision of the nodes in the system with the help of certificate and the thresholds limit.

The idea is that a device manages a network in which it stores information about the devices it meets: its network of trusted friends contains trusted devices, and its network of foes contains Malicious Devices. By reasoning on this network, the device is then able to determine whether an unknown individual is carrying out a sybil attack or not. The work will also evaluate the extent to which the proposed approach reduces the number of interactions with sybil attackers and consequently enables collaborative applications. The work will achieve this using real mobility and social network data. It also access computational efficiency.

IV. ALGORITHM

The algorithm is as follows:

```
Initiate SybilDecline {
  GetUserRegister==Social Network Service
  (U1, U2, U3,...Un);
  //Register Users with the System
  Each User Generates data;
  //User Request for Link
  Check_Users_Trust=>1.0;
  Identity==Trusted
  Trust_Certificate_Flag==1; //Authentic Node
  Message than allow operations
  For Each Node in Social Network Monitor the
  Behaviour against the Threshold Continuously
  ActivityMonitorFlag==0; //Initial monitoring
  of the System is Null
  Monitor User Behaviour==U1.TrackActivity
  (); // Start Activity Tracking for Selected User
  ActivityMonitorFlag==1;// Set Flag to 1 from 0
  Track Activity (Friend, Photo, Post, Profile
  Edit, Friend Request, Chat)
  {
  Assign Trust Value to Login, Session Time,
  File Sharing, Upload, Download, System for
  User;
  Monitor Counter (Login, Status, EditProfile,
  VistiProfile, ImageUpload, EditPass);
  GenerateGraph (Login);
  GenerateGraph (Status);
  GenerateGraph (EditProfile);
  GenerateTable (VistiProfile);
  GenerateTable (ImageUpload);
  GenerateTable (EditPass);
  MonitoringSystemCheck (CPU Load, RAM
  Utilization);
  Trust= $\sum_{i=1 \text{ to } n} \text{Count} * \text{Trust Weights}$ ;
```

```

U1TrustVerification=U1.SybilTrustModel (p1,
p2, p3, p4, p5, p6);
}
SybilTrustModel (p1, p2, p3, p4, p5,
p6)//Parameter for Each Users Activity
{
Trust Threshold==0.5;
If (Trust < -0.5)
Logout==True;
If (Trust<=-0.9)
UserDelete==True;
}
Reset ActivityMonitorFlag==0;
Trust_Certificate_Flag==0;
Exit;
}

```

V. EXPECTED OUTCOME

1. It is helpful to find sybil attacks thus making the system secure.
2. Easy detection using Trusted and Threshold values.
3. Trust value verification allows only authentic user to further forward the request.
4. Business intelligence applications more likely to be motivated after sybil removal.
5. It is used to find fake user identities.

TABLE I. User Activity Based Sybil Monitoring

| Sr. No | User Name | Number of Login (Count) | Activity Performed (Count) | | | | Output Interpretation (Sybil) [Yes/No/NA] |
|--------|-----------|-------------------------|----------------------------|---------------|---------------|--------------|---|
| | | | Profile Visit | Status Update | Password Edit | Image Upload | |
| 1 | aaa | 358 | 38 | 52 | 14 | 8 | Yes |
| 2 | bbb | 295 | 9 | 8 | 12 | 0 | Yes |
| 3 | ccc | 80 | 12 | 5 | 17 | 0 | Yes |
| 4 | ddd | 20 | 3 | 3 | 65 | 0 | Yes |
| 5 | yyy | 1 | 0 | 0 | 0 | 0 | No |
| 6 | pqr | 0 | 1 | 3 | 20 | 0 | No |

TABLE II. Users Trust and Decision Table

| Sr. No | User | Trust Values | | | | | | Decision (If Trust Val<=Min Threshold) [Action] |
|--------|------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---|
| | | TV1 After Activity 1 | TV2 After Activity 2 | TV3 After Activity 3 | TV4 After Activity 4 | TV5 After Activity 5 | TV6 After Activity 6 | |
| 1 | U1 | 0.4 | 0.7 | 0.75 | 0.52 | 0.84 | 0.96 | Normal |
| 2 | U2 | 0.4 | 0.52 | 0.63 | 0.71 | 0.84 | 0.9 | Normal |
| 3 | U3 | 0.3 | 0.2 | 0.4 | 0.7 | 0.8 | 0.6 | Normal |
| 4 | U4 | 0.4 | 0.5 | 0.48 | 0.41 | 0.43 | 0.3 | Normal |
| 5 | U5 | 0.4 | 0.25 | 0.18 | -0.2 | -0.31 | -0.52 | Sybil |
| 6 | U6 | 0.8 | 0.5 | 0.61 | 0.22 | 0.06 | -0.18 | Sybil |

VI. APPLICATION DOMAIN

Various application domains of this work include the following:

- Mobile Networks
- Auditing
- Cash Economies
- Reputation Systems
- Symmetric Reputation
- Asymmetric Reputation Systems

VII. EXPERIMENTAL EVALUATION

The goal of the proposed approach is to reject sybil's and accept honest people. To ascertain the effectiveness of proposed work evaluation factors give correct directions. The

user with higher statistical records or their activities and the trust values commonly derived the sybil behavior in the system. The above table I shows the increasing order of activity statics. According to the activities the users which are operating more on with different time based logins and the activities like status and password are updated very frequently then the user is considered as Sybil and the interpretation is given as YES. If the user is having less values then the output will be NO. If the statics was not able to derive some relation and results than the table II is used. It can also be used for confronting the Sybil detected by analytical evaluation.

The trust computation will depends on the user activities categorized in constructive (+ve) and destructive (-ve). All the values are assigned some weights for their respective operations performed by the user. On the basis of that the trust is evaluated and the final action is taken out as a decision of the approach. Once the system directs the user out from the system it means the user is a Sybil user. After being performing those activities the values are added or subtracted

from the composite trust. If the value of trust is lesser than a predefined given threshold then the user is considered as a sybil node and gets moved out of the system. In this way a higher protection level can be achieved here with less computation and overhead. Later on re login or permanent removal can be applied according to the administrative policies.

The table I shows different activities performed by the user at different instance of time and their respective trust value assigned according the final trust value is then used to find the suggested Sybil or non Sybil nodes with the help formula IF (trust value \leq MinThreshold) Action as shown in the column Decision .The Minimum threshold value is set to 0.5 accordingly decisions are taken for users in the system.

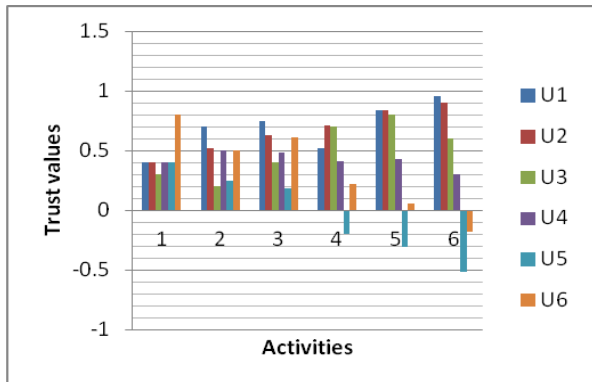


Figure 1: Trust comparison between different users for sybil detection.

From the graph in Figure 1 we can conclude that users U1,U2,U3,U4 are trusted user with their trust count 0.96,0.9,0.6,0.3 and the suggested Sybil users are U5 and U6 with negative graph and the value of trust count -0.52,-0.18 respectively. Separately analyzing the activities of user1 and user2 for a particular time instance is shown in the Figure 2.

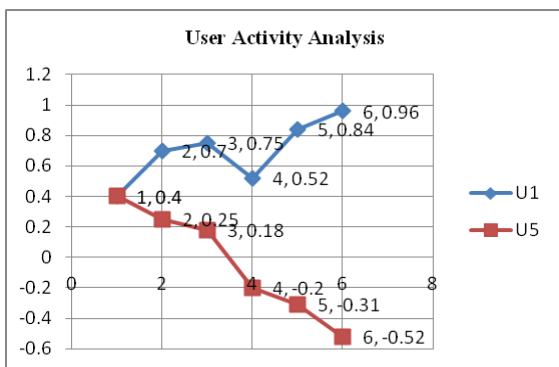


Figure 2. Trust values of user with respect to time

The graph in the figure 2 shows the activities of two users and their associated trust values .The final trust value helps to find out the whether the node is Sybil or honest node. Both the user start at the value 0.4 allocated to them during the registration then the values increases or decreases according to the activities and its associated trust value finally at the end of the first time instance the user1(U1) is labeled as honest node or normal node and the user two(U5) is labeled as (Suggested) Sybil node.

VIII. RESOURCE CONSTRAINED ANALYSIS

The system operations were executed for two conditional states i.e. before execution (Figure 3a) and after execution (Figure 3b) and studied using a tool jconsole [19]. By the above results, it is found that the system resource constraints are improving their values after execution, which shows that the execution is lightweight. The impact on the various parameters such as heap memory usage, number of threads, number of classes and CPU utilization is not majorly changes even with the presence of our tool of sybil detection. While analyzing the values it is clear that the practical implementation suggested for our proposed approach is lightweight in nature and serve its purpose with very less impact on its other primary processes. Heap memory is used to support the memory requirements of the developed system. By the graph it is analyzed that the system is having reduced memory requirements in terms of heap. Also the consumption of resources demanding memory is also optimized here. It also support dynamic allocation from which the memory is been managed between underutilized and over utilized processes. The number of classes supports the phenomenon of cardinality which gives that for execution the class count must be higher. It separates the different types of data into various classes of similar type of attributes or characteristics. While carefully watching the graph we found that after execution the class counts is increased which means that the supported classes was increased to have better resource utilization. The number of class count and the original class count should be in minimal range. As seen by the graph in the figure 3a and figure 3b number of threads created by the server to execute the task is similar to that before execution. Thus the work can be able to serve its results without any sudden increase in thread count. Also the similar thread count shows that the developed prototype is having constant parallelism which gives effective analysis and utilization of resources. CPU Usage gives the amount of computer usage for processing resources, or the amount of work done by the CPU. It varies on the amount and type of managed task done .In the figure CPU usage graph is shown between 0.0% to 1.0% before the system was started the graph shown an average of 0.1% (approximate) and after running the system the average CPU utilization was maintained from 0.8% to finally 0.1% (approximate) this shows that the system is a light weighted. CPU utilization may be used to gauge system performance. That is with a heavy load with only a few running programs may indicate insufficient CPU power support ,or running programs hidden by system monitor ,a high indicated presents of unwanted element in the system.

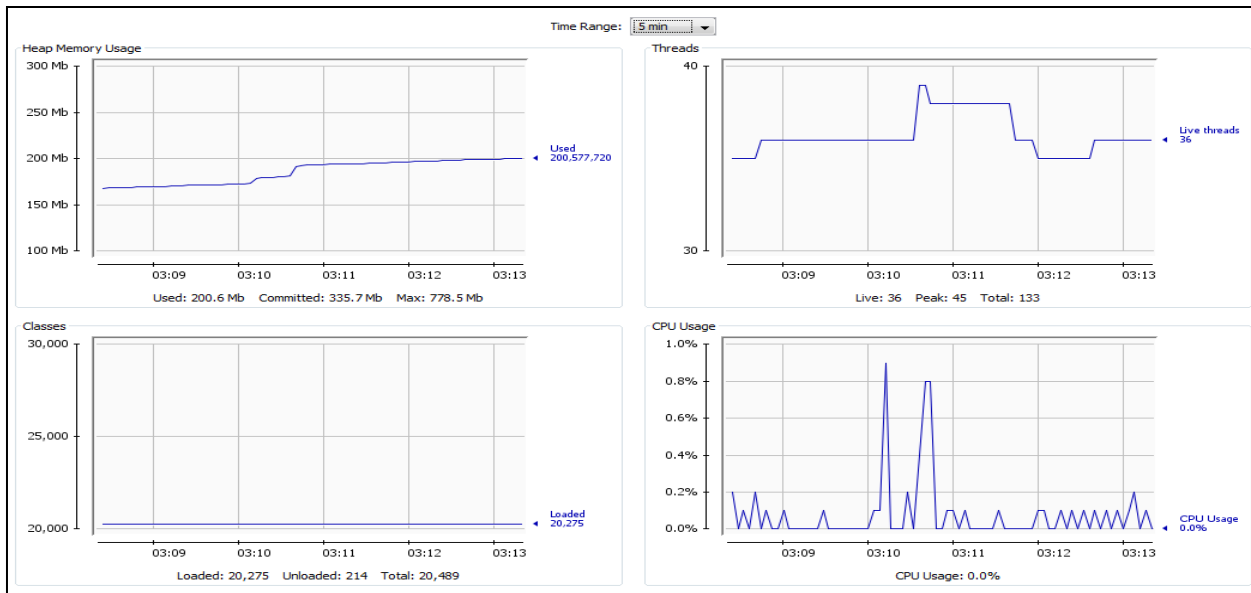


Figure 3a : Before execution (Parameters Heap Memory Usage, Threads, Classes and CPU Usage)

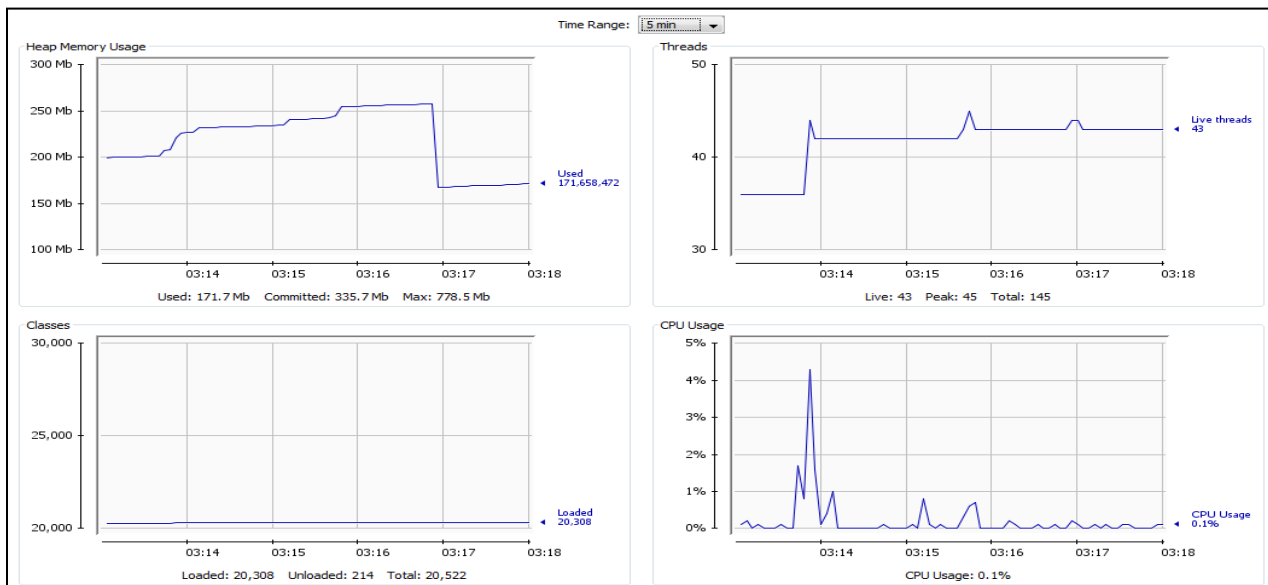


Figure 3b: After execution (Parameters Heap Memory Usage, Threads, Classes and CPU Usage)

IX. CONCLUSION

Users are abruptly increasing in the social media causing a wide variation in the behavior and frequency of content access. As this request is very large in numbers, their verification and identification of honest or forge identities is moving towards critical zone. Many of the users are interested in such activities which required some fake identity creations which might be made for some disruptions to actual user's data. Thus to provide a secure mechanisms which handles such request and identities and let them secure from attackers is a emergent need of the market. Such forged identities are called as Sybil and their prevention mechanism is called Sybil attack detections. In the course of time, there have been much a large number of strategies suggested to conquer such problems of fake

user conduct still there stay a few issues unaddressed. This work gives a novel Sybil attack identification and evacuation mechanism based on the behavioral conduct analysis, trust and threshold calculation. Sybil attack is most crucial in distributed decentralized systems like social networks. Since too much personal and sensitive information is posted openly in social networks, the threats posed by Sybil attacks are very serious. Here we have discussed the need for social network security and the effects of Sybil attacks in social networks, classification of Sybil attacks to understand about the sources of the attack, 11 examples of systems vulnerable to Sybil attacks and a summary of the current trends in Sybil defenses. The Sybil defenses are classified according to the attacker's capability and behavior attributes of Sybil user's. After prelim experiments the orientation and directions of works

gives a positive outcomes and assures its applicability in near future.

ACKNOWLEDGEMENT

Our thanks to all the anonymous referral for their comments which paved a way for strengthening this paper and also like to thank teachers and friends for their discussions in the respective domain.

REFERENCES

- [1] Wei, Fengyuan Xu, Chiu C. Tan and Qun Li, "SybilDefender: Defend Against Sybil Attacks in Large Social Networks", in IEEE Transaction, Dec 2012.
- [2] Bimal Viswanath and Ansley Post, "An Analysis of Social Network-Based Sybil Defenses", in ACM Special Issue on SIGCOMM-10, doi: 978-1-4503-0201-2/10/08, Nov 2010.
- [3] Zhi Yang, Christo Wilson and Xiao Wang, "Uncovering Social Network Sybils in the Wild", in ACM Special Issue on IMC-11, doi: 978-1-4503-1013-0/11/11, Nov 2011.
- [4] Abedelaziz Mohaisen, Huy Tran, Nicholas Hopper and Yongdae Kim, "On the Mixing Time of Directed Social Graphs and Security Implications", in ACM Special Issue on ASIACCS-'12, doi: 978-1-4503-0564-8/11/03, May 2012.
- [5] Bimal Viswanath, Mainack Mondal, Allen Clement and Peter Druschel, "Exploring the design space of social network-based Sybil defenses", in IEEE Transaction, doi: 978-1-4673-0298-2/12, 2012.
- [6] Haifeng Yu, Michael Kaminsky, Phillip B. Gibbons and Abraham Flaxman, "SybilGuard: Defending against Sybil Attacks via Social Networks", in ACM Special Issue on SIGCOMM-06, doi: 1595933085/06/0009, Sep 2006.
- [7] Zhuhua Cai and Christopher Jermaine, "The Latent Community Model for Detecting Sybil Attacks in Social Networks", in ACM Special Issue on VLDB Endowment, Sep-2011.
- [8] Frank Li, Prateek Mittal, Matthew Caesar and Nikita Borisov, "SybilControl: Practical Sybil Defense with Computational Puzzles", in ACM Special Issue on STC-12, doi: 978-1-4503-1662-0/12/10, Oct-2012.
- [9] Haifeng Yu, Phillip B. Gibbons, Michael Kaminsky and Feng Xiao, "SybilLimit: A Near-Optimal Social Network Defense against Sybil Attacks", in IEEE/ACM Transaction, ISSN: 1063-6692, doi: 10.1109/TNET.2009.2034047, 2009.
- [10] Yazan Boshmaf, "A Quick Survey of Social Network-based Sybil Defenses", in article at University of British Columbia, Vancouver, Canada.
- [11] Lu Shi, Shucheng Yu, Wenjing Louy and Y. Thomas Hou, "SybilShield: An Agent-Aided Social Network-Based Sybil Defense among Multiple Communities", in Proceedings of IEEE Infocomm-13, doi: 978-1-4673-5946-7/13, 2013.
- [12] Renuga Devi R and M. Hemalatha, "Sybil Identification in Social Networks Using SICT and SICTF Algorithms with Improved KD-Tree", in Journal of Theoretical and Applied Information Technology (JATIT), ISSN: 1992-8645, Vol. 56 No.2, Oct-2013.
- [13] Philip W. L. Fong, "Preventing Sybil Attacks by Privilege Attenuation: A Design Principle for Social Network Systems", in IEEE Symposium on Security and Privacy, ISSN: 1081-6011/11, doi: 10.1109/SP.2011.16, 2011.
- [14] Bastian Konings, David Piendl, Florian Schaub, and Michael Weber, "PrivacyJudge: Effective Privacy Controls for Online Published Information", in 3rd IEEE International Conference on Privacy, Security, Risk and Trust (PASSAT 2011), at: <http://dx.doi.org/10.1109/PASSAT/SocialCom.2011.86>, 2011.
- [15] Bo Zhu, Sushil Jajodia and Mohan S. Kankanhalli, "Building trust in peer-to-peer systems: a review", in Int. J. Security and Networks, Vol. 1, Nos. 1/2, 2006. Pp 103-112.
- [16] Peng Zhou, Xiapu Luo, Ang Chen, and Rocky K. C. Chang, "STor: Social Network based Anonymous Communication in Tor", in Hong Kong Polytechnic University, archive arXiv: 1110.5794v6, 2013.
- [17] Thibault Cholez, Isabelle Christment, Olivier Festor and Guillaume Doyen, "Detection and mitigation of localized attacks in a widely deployed P2P network", in Springer Peer-to-Peer Networking and Applications, Volume 6, Issue 2, DOI 10.1007/s12083-012-0137-7, 2012. pp 155-174.
- [18] Deepti Sharma and Dr. Sanjay Thakur, "Sybil-Deliver: A Survey on Novel Trusted Identity and Threshold Based Path Rank for Sybil Attack Identification in Social Network" in IJARCSSE, ISSN: 2277:128X, Volume 4, Issue 6, June 2014.
- [19] Using Jconsole to Monitor applications, "<http://www.oracle.com/technetwork/articles/java/jconsole-1564139.html>".

Cost and Performance Based Comparative Study of Top Cloud Service Providers

Irfan Bari^{#1}, Salaser Babu^{*2}, Muhammad Munwar Iqbal³, Yasir Saleem⁴, Zahid Ali Masood

^{1, 2, 3} Department of Computer Science, University of Engineering and Technology, Taxila

^{3, 4} Department of Computer Science & Engineering, University of Engineering and Technology, Lahore

Abstract— Recent boom in the cloud computing industry has caused a shift in the information technology industry and has affected the way how information and data is stored and shared among the enterprise. The advent of social applications also demands the availability of resources that can be shared among the others. Cloud based architecture has made it possible for enterprises to utilize the computation power that was not available in the past. This paper takes a look and compares the top available service providers on the basis of the cost for each computing model as well takes a look at the performance by measuring the response time. It is observed that at all these service providers and elaborated the comparison of all based on their available architectures.

Keywords— Comparison of Cloud Operators, Cloud Computing, Azure, RackSpace, Amazon Web Services

I. INTRODUCTION

As every new technology has to go through some initial stages in order to achieve some maturity level, Cloud computing is also one of the struggling technologies trying to get out of its infancy stage. Cloud computing has introduced the new way of computation, this have changed the domain of physical infrastructural setup, instead unpacking, installation and setup of servers, user can get everything up and running in minutes and there are no limits to hardware, everything is managed virtually.[1] The term cloud computing or simply cloud is a new way of data storage and remote access to online available resources in order to achieve coherent and economy of scale utility. To avail the advantages of cloud computing, most of the today's companies are being shifted to this emerging paradigm. The shift from traditional CAPEX based approach to OPEX based model is a basic definition of the term "moving to cloud". Recent researches conducted on cloud computing reveals that there are several definitions with different aspects were proposed by most of the researchers throughout the course of cloud's development. Maricela-Georgiana Avram (Olaru) says "With all of the hype around cloud computing, and multiple definitions of cloud computing, it is difficult to discern exactly what constitutes "cloud computing." [2]. Cloud computing has provided a low cost infrastructural access that was never available before [3]. A company maintained data centre is costly with respect to

operation and maintenance. Cloud is an alternate source to achieve the desired computation power with the option to enhance at any time [4].

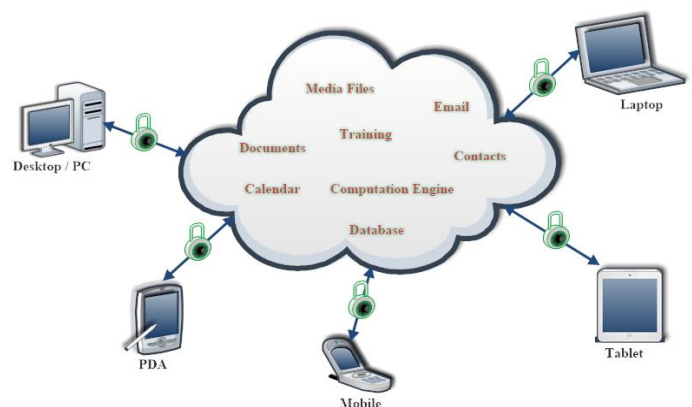


Figure 1 Basic Cloud Infrastructure, there is no boundary of nodes or hardware to interact with the cloud as cloud architecture is based on services and every device can communicate and share data via authenticated channel

II. ADVANTAGES AND CHALLENGES OF CLOUD COMPUTING

A. Economical

Cloud computing is highly beneficial for the enterprise to novice companies as it provides an economical and least expensive way to the resources which may be unbearable in most of the circumstances. Within a limited set of budget the organization get the access to the state of art hardware and infrastructure [5]. Maricela-Georgiana Avram (Olaru) says "It can provide an almost immediate access to hardware resources, with no upfront capital investments for users" [2]

B. Unleash the Innovation

The recent shift to cloud has opened a new breed of computational power which makes the availability of exceptional resources that can be utilize to innovate the new class of application software which was never possible before the introduction of cloud [6]. For example Facebook, DropBox, Google Drive etc are one the few examples which are serving millions of users.

C. Security and Privacy

One of the biggest challenges for cloud systems is security. As this is under the process of evolution and unlimited number of configuration, there is a need to have a unified infrastructure that can be used to ensure the integrity of data[7]. There are multiple security models proposed which have different level of implementations and this again causes confusion among the enterprises which are reluctant to shift to cloud computing

D. Infrastructural shift

As the hardware and architecture of cloud is different from the private server infrastructure, thus this requires a comprehensive shift in process how the computational platform works.[8] This causes acquisition of new skills as well as development of new applications to accommodate the shift. As a concern for enterprise, this sometimes makes it rather difficult if not impossible to shift as they have a large set of complex operational and computational applications which are the core for their operational activities

III. RELATED WORK

After reviewing the study done in the field of comparative study of cloud service providers, it is clear that there have been a fewer number of studies reported on comparison of cloud computing systems.

Buyya et al.[9] compared the selected cloud service providers on the basis of market orientation. He tried to analyse the current demands of the market and consumers and tried to analyse the risks involved in it. Zhang et al. [10]

worked on the generic approach and compared four selected cloud service providers on the basis of diversified services. These services are abstracted from consumer but to provide the desired functionality. After his review it was resulted that in order to get all the functionality, one cannot rely on a single cloud service provider. Voorsluys et al.[11] look as the attributes of cloud service providers on the basis of Infrastructure-As-A-Service(IAAS) and Platform-As-A-Service (PAAS).

In some other studies another approach is utilized based on taxonomic attributes. In this field of study, Hofer et al. [12] introduced tree based taxonomic structure for cloud service providers. In this study, cloud service providers were taxonomically characterized on the basis of characteristics / attributes like service model, license type, cost model, supported languages and operating systems, development tools and virtualization mechanism etc. Another work in this regard was done by Rimal et al.[13] who introduced taxonomy based features that are provided by cloud service providers. The attributes selected for the comparison were provider's architecture, virtualization mechanism, services provided, mechanisms for load balancing, fault tolerance, interoperability and security, and support for software and programming languages etc. In some of the studies, quantitative comparisons have been draw among different providers, Cloud Comp, for instance, is a framework proposed to provide performance comparison of various providers.

Table 1: Summary of Various Authors Work with respect to cloud service providers

| | Objective | Providers | Features | Type | Results |
|---------------------|---|---|---|-------------|--|
| Zhang et al. | Analysis of features of selected cloud service providers | EC2 Azure AppEngine | Services Applications virtualizations scalability | Qualitative | Current service providers have a variety of features and the user cannot sustain utilizing only one source |
| Buyya et al. | Market based cloud services comparison | EC2 AppEngine Live Mesh sunnetwork GRIDS | Services virtualization QoS control | Qualitative | Cloud service providers don't comply with prevailing market environment |
| Rimal et al. | Introduction of taxonomy to classsify cloud service providers | AWS GoGrid FlexiScale Mosso AppEngine GigaSoaces Azure RightScale SunCloud Salesforce | Virtualization Services scalability reliability interoperability security | Taxonomy | Sharing of resources must be compensated based on the contribution |
| Hofer et al. | Introduction of taxonomy to classify cloud service providers | Azure AppEngine | Services licensing user type payment QoS security | Taxonomy | Standard operational procedures are or be adapted for resolving the issues like interoperability |
| Li et al. | Framework to compare multiple attributes of cloud service providers | AWS Rackspace | Computation storage performance | Quantities | Variations exists in performance in different aspects |

Table 1 summarizes different research options reported with reference to the topic in study. All these studies are classified based on the type of comparison made by the authors. As the number of providers currently available are very large (more than 80 according to [15]), therefore, cannot say that these studies cover all of cloud based service providers and neither it represents the outlook of all of them. Due to the large number of cloud providers currently available. Therefore it is always needed to get a comparative analysis which represents the current running scenario which is helpful for everyone to select the right service provider.

IV. AVAILABLE CLOUD SERVICE PROVIDER

With the rapid advancement in the field of cloud computing business, the number of cloud service providers is increasing at an enormous pace.

Customers have so many choices that were never available before. Now users face so many technical and infrastructural options that were never available before.[16]

Keeping in view recent trends in IaaS, below is a comparison among top 5 cloud companies on the basis of below mentioned parameters

Table 2 Top Five Cloud Service Providers based in the surveys conducted and customer base

| | Service Providers | | | | |
|----------------------------|----------------------------------|----------------------------|--------------------------------------|------------------------------|------------------------------|
| | AWS | Azure | Rackspace | Hp Cloud | IBM Softlayer |
| YEAR | 2006 | 2010 | 2010 | 2012 | 2005 |
| No. of Data Centers | 11 | 16 | 6 | | 7 |
| Supported OS | Microsoft Windows Linux FreeBSD | Microsoft Windows Linux | Microsoft Windows Linux | Microsoft Windows Linux | Microsoft Windows Linux |
| Supported Language | Java. PHP. Python. Ruby. ASP.NET | PHP ASP.NET Node.js Python | C#.NET, Python, PHP, Java, and Ruby. | Ruby PHP Java Node.js Python | ASP.NET Perl PHP Python Ruby |

V. SELECTED CLOUD SERVICE PROVIDERS

The top most cloud service providers are chosen based on their market share and number of customers.

A. Amazon Web Services

Amazon Web Services has been one of the famous cloud hosting platforms, it was introduced in 2006 as a collection of remote computing services. AWS provides services nearly in

all domain of cloud computation ranging from computation to database and content delivery networks.

Its use is not limited to any particular domain but in fact AWS can find AWS providing service to nearly every aspect. Some of these are

- 1) *Social Networks*: Sound Cloud, Foursquare, Shazam, Pinterest, Netflix, Imdb, 9gag, Oasis, Reddit, Transfer, Yelp, Hasoffers
- 2) *Document/Educational*: Scribd, Coursera, University of California Berkeley, University of Chicago, University of Notre Dame
- 3) *Organizations / Brands*: Siemens, Ericsson, Lionsgate, Novartis, Pfizer, Qantas, Sap, Schneider Electric, Adobe, Trend Micro, Uni Lever, Vodafone, Xiaomi, Zyxel, Lamborghini, Nokia, HTC, Property Guru
- 4) *Media*: Newsweek, The Guardian, Score Media, ScribbleLive, United Daily News

B. Windows Azure

Windows Azure was released in 2010 as a cloud computational platform and infrastructure. Azure has nearly same services domain as of AWS with difference that it provides some additional components that are proprietary of Microsoft like visual studio online, analytics and insight etc.

Some of the Azure services users are

- 1) *Social Network*: Accuweather
- 2) *Documents / Education*: Marquette University, Federal University of Rio Grande do Sul, Carnegie Mellon University, Royal Holloway, University of London, Tokyo University of Technology, University of California, University of Rome Tor Vergata
- 3) *Organizations / Brands*: Telenor, SK Telecom, ICICI Lombard, Aston Martin, Dell, Mazda, Samsung, SAAB, 3M, Xerox, Trek, XBOX, Microsoft
- 4) *Media*: NBC, iStreamPlanet, MagV, NetGem, News360.

C. Rackspace

Rackspace introduced as utility computing offering in 2006. It was renamed to Rackspace Cloud in 2009. Rackspace was ranked 29th in Fortune 100 for best companies to work for in USA in 2014. Rackspace provides computation engine as well as database and content delivery network services.

Here are some of the famous clients of RackSpace

- 1) *Social Networks*: Shaadi.com, Sprout Social, Pearl.com, Didlr, Scout Mob
- 2) *Document / Educational*: Ethica Institute of Islamic Finance, After College, Institute of Engineering and technology UK
- 3) *Organizations / Brands*: Domino Pizza, Gorilla, Rokkan, Simon Property Group, X Commerce, Mazda

North America, Coastal , Metro Trains Melbourne, GeoData Plus

- 4) *Media*: Blue Bell Group, HUGE

VI. COMPARATIVE STUDY

The conducted study based on the following two parameters

1. Cost
2. Performance

A. Cost

The cost models have been divided into three categories based on the computation power. These three categories are

1. Small scale computation

Small scale computing class consist of most specific computational powers with normally one CPUs and memory ranging up to 4GB. After this classification found that there are 3 types of hardware architectural configurations which fall under this category. This include

- a. T2 Micro with 1CPU and 1GB RAM
- b. T2 Small with 1CPU and 2GB RAM
- c. T2 Medium with 1CPU and 4GB of RAM

2. Medium scale computation

Medium Scale computation class include number of CPUs ranging from 1to 13 CPUs while the memory limit is from 1.7GB to 15GB. Medium Scale classification have further two sub divisions based on the hardware configuration. These are divided into M3 and M1 categories.

M3 includes

- a. M3 Medium with 1CPU and 3.75GB RAM
- b. M3 Large with 2CPUs and 7.5GB RAM
- c. M3 xLarge with 13 CPUs and 15Gb RAM

While M1 includes

- a. M1 Medium with 2CPUs and 3.75Gb RAM
- b. M1 Large with 4CPUs and 7.5GB RAM
- c. M1 xLarge with 8CPUs and 15GB RAM

3. Large scale computation

Large scale computing class is based on the number of CPUs ranging from 14 to 88 and collective memory to ranging from 7GB to 60.5GB

With criteria based on the above mentioned classification there are two categories one is about clusters and second is about high end CPUs

In cluster have

- a. Cc2.8xLarge with 88CPUs and 60.5GB RAM
- b. Cc g2.2xLarge with 26 CPUs and 15GB RAM
- c. Cg 1.4xLarge with 33.5CPUs and 22GB RAM

While High End category include

- a. C3.xLarge with 14CPUs and 7GB of RAM
- b. C3.2xLarge with 28CPUs and 15GB RAM
- c. C3.4xLarge with 55 CPUs and 30GB RAM

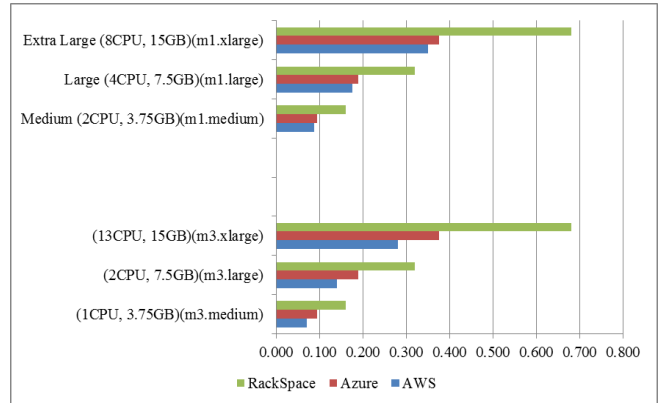


Figure 2 Comparison of medium scale computing of all three cloud service providers based on CPU numbers and Memory size

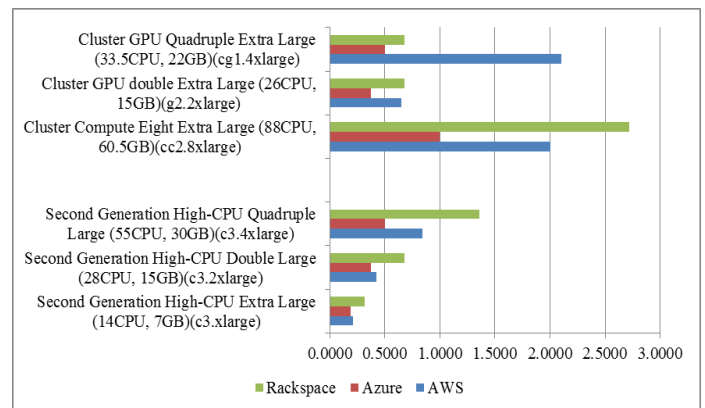


Figure 3 Comparison of large scale computing of all three cloud service providers, mostly containing the high end cluster based architecture

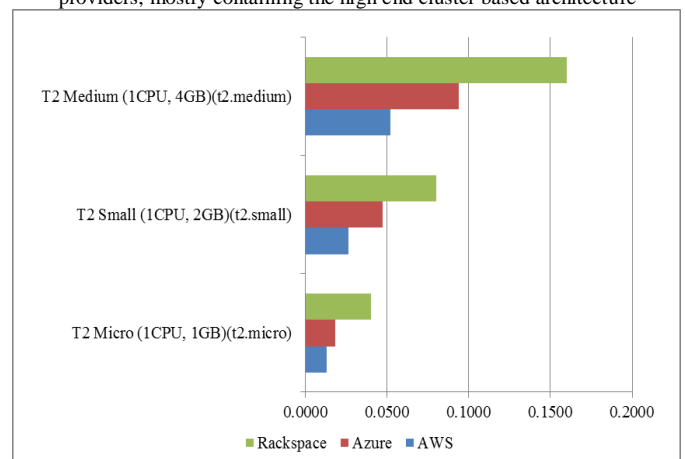


Figure 4 Comparison of small scale computing of all three cloud service providers

B. Performance

In the performance perspective, experiment is conducted on the based on response time.

For this purpose, used PHP SDKs of all three cloud service providers. To measure the response, a basic application with login authentication was developed and the response time was measured as the difference in time from the point of call being sent to the server and response received from the server.

The hardware and software used in this experiment was HP 4530s probook with core i3 2310 2.1Ghz, 4 GB RAM, 320 GB Toshiba 7200 RPM hard disk, Windows 8.1. Internet Connection Bandwidth 2MB WAMP server version 2.2 with PHP version 5.3.13 and APACHE version 2.2.22.

To reduce the error rate and to get average, performed the experiments for 10 times on different time and date to get the results. On the basis of these results are derived the following graph.

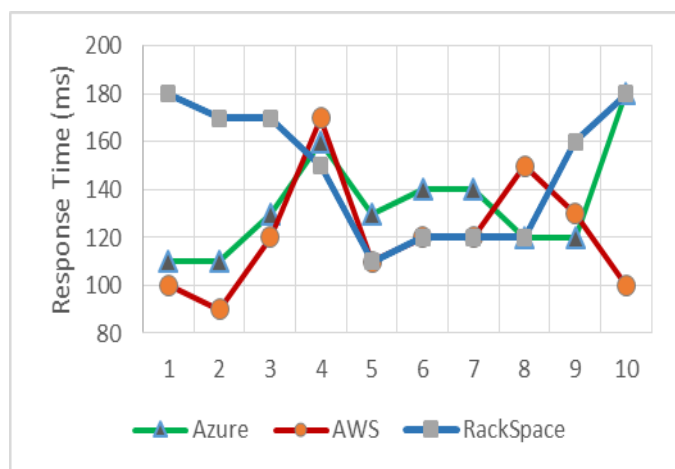


Figure 5 Comparison of medium scale computing of all three cloud service providers based on the response time in milliseconds compared with the average number of attempts made

VII. CONCLUSIONS AND FUTURE WORK

With reference to figure 2, 3 and 4 can clearly elaborate that in case of small and medium scale computing models, AWS has best price options available. While in case of large scale computing, azure has an edge. While RackSpace is most costly in all cases.

So if analyze all three cloud service providers than AWS has the edge and thus it is the most cost effective solution, in the current circumstances. But, it must have to keep in mind the geographical location and the proximity of the cloud and bandwidth of internet that all has a major role in this conclusion as well as the number of sites which are hosted at the particular cloud also affects the availability of paths to the endpoint. Because as the access routes are refreshed more in case of network traffic from a particular data center.[17]

Amazon also has an edge on Azure and RackSpace because it has been in operation from more time than the other two.

AWS has cloud operations and services in much mature levels. And also the number of data centers and major companies like Siemens, NOKIA, Adobe etc and social network infrastructure like Sound Cloud, Pinterest makes it more appropriate.

Similarly if look at the security models of all these cloud service providers, all of them have an exceptional model to work this out. In case of Amazon, the classical NIST cloud model followed with the encryption and access control option are deployed to provide a secure and reliable computation experience. While in case of Azure, Service Management Portal / API section provides the secure access to the data to owners and developer while users are provided access to the resources using fabric layer and access control management. Rackspace security model is based on Dr. W. Edwards Deming's Plan-DoCheck-Act management method which uses plan, do, check and act steps. So far our study is based on measurement of response time for the chosen cloud service providers. Our future work will be including the comparative study of computational capabilities of all these three service providers. For this developing a different application that will be used to measure the requested parameters

REFERENCES

- [1]. Tebaa, M., and Hajji, S.E.L.: 'From Single to Multi-clouds Computing Privacy and Fault Tolerance', IERI Procedia, 2014, 10, (0), pp. 112-118
- [2]. Avram, M.G.: 'Advantages and Challenges of Adopting Cloud Computing from an Enterprise Perspective', Procedia Technology, 2014, 12, (0), pp. 529-534
- [3]. Rabai, L.B.A., Jouini, M., Aissa, A.B., and Mili, A.: 'A cybersecurity model in cloud computing environments', Journal of King Saud University - Computer and Information Sciences, 2013, 25, (1), pp. 63-75
- [4]. Dianat, O., and Orgun, M.A.: 'Representing and Reasoning about Utilization of Cloud Computing as Bayesian Games with Epistemic Logic', Procedia Computer Science, 2013, 19, (0), pp. 40-47
- [5]. Ercan, T.: 'Effective use of cloud computing in educational institutions', Procedia - Social and Behavioral Sciences, 2010, 2, (2), pp. 938-942
- [6]. Shahzad, F.: 'State-of-the-art Survey on Cloud Computing Security Challenges, Approaches and Solutions', Procedia Computer Science, 2014, 37, (0), pp. 357-362
- [7]. Che, J., Duan, Y., Zhang, T., and Fan, J.: 'Study on the Security Models and Strategies of Cloud Computing', Procedia Engineering, 2011, 23, (0), pp. 586-593
- [8]. Cheng, F.-C., and Lai, W.-H.: 'The Impact of Cloud Computing Technology on Legal Infrastructure within Internet—Focusing on the Protection of Information Privacy', Procedia Engineering, 2012, 29, (0), pp. 241-251
- [9]. Buyya, R., Yeo, C.S., Venugopal, S., Broberg, J., and Brandic, I.: 'Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility', Future Generation computer systems, 2009, 25, (6), pp. 599-616
- [10]. Zhang, Q., Cheng, L., and Boutaba, R.: 'Cloud computing: state-of-the-art and research challenges', Journal of internet services and applications, 2010, 1, (1), pp. 7-18
- [11]. Voorsluys, W., Broberg, J., and Buyya, R.: 'Introduction to cloud computing', Cloud computing: Principles and paradigms, 2011, pp. 1-44

- [12]. Höfer, C., and Karagiannis, G.: 'Cloud computing services: taxonomy and comparison', *Journal of Internet Services and Applications*, 2011, 2, (2), pp. 81-94
- [13]. Rimal, B.P., Choi, E., and Lumb, I.: 'A taxonomy and survey of cloud computing systems', in Editor (Ed.)^(Eds.): 'Book A taxonomy and survey of cloud computing systems' (Ieee, 2009, edn.), pp. 44-51
- [14]. Ostermann, S., Iosup, A., Yigitbasi, N., Prodan, R., Fahringer, T., and Epema, D.: 'A performance analysis of EC2 cloud computing services for scientific computing': 'Cloud computing' (Springer, 2010), pp. 115-131
- [15]. 'Compare cloud service providers <http://cloud-computing.softwareinsider.com/>'
- [16]. Wang, M., Wang, G., Tian, J., Zhang, H., and Zhang, Y.: 'An Accurate and Multi-faceted Reputation Scheme for Cloud Computing', *Procedia Computer Science*, 2014, 34, (0), pp. 466-473
- [17]. Barceló, J. M., Nieto-Hipólito, J. I., and García-Vidal, J.: 'Study of Internet autonomous system interconnectivity from BGP routing tables', *Computer Networks*, 2004, 45, (3), pp. 333-344

IJCSIS REVIEWERS' LIST

Assist Prof (Dr.) M. Emre Celebi, Louisiana State University in Shreveport, USA
Dr. Lam Hong Lee, Universiti Tunku Abdul Rahman, Malaysia
Dr. Shimon K. Modi, Director of Research BSPA Labs, Purdue University, USA
Dr. Jianguo Ding, Norwegian University of Science and Technology (NTNU), Norway
Assoc. Prof. N. Jaisankar, VIT University, Vellore, Tamilnadu, India
Dr. Amogh Kavimandan, The Mathworks Inc., USA
Dr. Ramasamy Mariappan, Vinayaka Missions University, India
Dr. Yong Li, School of Electronic and Information Engineering, Beijing Jiaotong University, P.R. China
Assist. Prof. Sugam Sharma, NIET, India / Iowa State University, USA
Dr. Jorge A. Ruiz-Vanoye, Universidad Autónoma del Estado de Morelos, Mexico
Dr. Neeraj Kumar, SMVD University, Katra (J&K), India
Dr Genge Bela, "Petru Maior" University of Targu Mures, Romania
Dr. Junjie Peng, Shanghai University, P. R. China
Dr. Ilhem LENGILIZ, HANA Group - CRISTAL Laboratory, Tunisia
Prof. Dr. Durgesh Kumar Mishra, Acropolis Institute of Technology and Research, Indore, MP, India
Dr. Jorge L. Hernández-Ardieta, University Carlos III of Madrid, Spain
Prof. Dr.C.Suresh Gnana Dhas, Anna University, India
Dr Li Fang, Nanyang Technological University, Singapore
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Dr. Siddhivinayak Kulkarni, University of Ballarat, Ballarat, Victoria, Australia
Dr. A. Arul Lawrence, Royal College of Engineering & Technology, India
Dr. Wongyos Keardsri, Chulalongkorn University, Bangkok, Thailand
Dr. Somesh Kumar Dewangan, CSVTU Bhilai (C.G.)/ Dimat Raipur, India
Dr. Hayder N. Jasem, University Putra Malaysia, Malaysia
Dr. A.V.Senthil Kumar, C. M. S. College of Science and Commerce, India
Dr. R. S. Karthik, C. M. S. College of Science and Commerce, India
Dr. P. Vasant, University Technology Petronas, Malaysia
Dr. Wong Kok Seng, Soongsil University, Seoul, South Korea
Dr. Praveen Ranjan Srivastava, BITS PILANI, India
Dr. Kong Sang Kelvin, Leong, The Hong Kong Polytechnic University, Hong Kong
Dr. Mohd Nazri Ismail, Universiti Kuala Lumpur, Malaysia
Dr. Rami J. Matarneh, Al-isra Private University, Amman, Jordan
Dr Ojesanmi Olusegun Ayodeji, Ajayi Crowther University, Oyo, Nigeria
Dr. Riktresh Srivastava, Skyline University, UAE
Dr. Oras F. Baker, UCSI University - Kuala Lumpur, Malaysia
Dr. Ahmed S. Ghiduk, Faculty of Science, Beni-Suef University, Egypt
and Department of Computer science, Taif University, Saudi Arabia
Dr. Tirthankar Gayen, IIT Kharagpur, India
Dr. Huei-Ru Tseng, National Chiao Tung University, Taiwan

Prof. Ning Xu, Wuhan University of Technology, China
Dr Mohammed Salem Binwahlan, Hadhramout University of Science and Technology, Yemen
& Universiti Teknologi Malaysia, Malaysia.
Dr. Aruna Ranganath, Bhoj Reddy Engineering College for Women, India
Dr. Hafeezullah Amin, Institute of Information Technology, KUST, Kohat, Pakistan
Prof. Syed S. Rizvi, University of Bridgeport, USA
Dr. Shahbaz Pervez Chattha, University of Engineering and Technology Taxila, Pakistan
Dr. Shishir Kumar, Jaypee University of Information Technology, Wakanaghat (HP), India
Dr. Shahid Mumtaz, Portugal Telecommunication, Instituto de Telecomunicações (IT) , Aveiro, Portugal
Dr. Rajesh K Shukla, Corporate Institute of Science & Technology Bhopal M P
Dr. Poonam Garg, Institute of Management Technology, India
Dr. S. Mehta, Inha University, Korea
Dr. Dilip Kumar S.M, Bangalore University, Bangalore
Prof. Malik Sikander Hayat Khiyal, Fatima Jinnah Women University, Rawalpindi, Pakistan
Dr. Virendra Gomase , Department of Bioinformatics, Padmashree Dr. D.Y. Patil University
Dr. Irraivan Elamvazuthi, University Technology PETRONAS, Malaysia
Dr. Saqib Saeed, University of Siegen, Germany
Dr. Pavan Kumar Gorakavi, IPMA-USA [YC]
Dr. Ahmed Nabih Zaki Rashed, Menoufia University, Egypt
Prof. Shishir K. Shandilya, Rukmani Devi Institute of Science & Technology, India
Dr. J. Komala Lakshmi, SNR Sons College, Computer Science, India
Dr. Muhammad Sohail, KUST, Pakistan
Dr. Manjaiah D.H, Mangalore University, India
Dr. S Santhosh Baboo, D.G.Vaishnav College, Chennai, India
Prof. Dr. Mokhtar Beldjehem, Sainte-Anne University, Halifax, NS, Canada
Dr. Deepak Laxmi Narasimha, University of Malaya, Malaysia
Prof. Dr. Arunkumar Thangavelu, Vellore Institute Of Technology, India
Dr. M. Azath, Anna University, India
Dr. Md. Rabiul Islam, Rajshahi University of Engineering & Technology (RUET), Bangladesh
Dr. Aos Alaa Zaidan Ansaef, Multimedia University, Malaysia
Dr Suresh Jain, Devi Ahilya University, Indore (MP) India,
Dr. Mohammed M. Kadhum, Universiti Utara Malaysia
Dr. Hanumanthappa. J. University of Mysore, India
Dr. Syed Ishtiaque Ahmed, Bangladesh University of Engineering and Technology (BUET)
Dr Akinola Solomon Olalekan, University of Ibadan, Ibadan, Nigeria
Dr. Santosh K. Pandey, The Institute of Chartered Accountants of India
Dr. P. Vasant, Power Control Optimization, Malaysia
Dr. Petr Ivankov, Automatika - S, Russian Federation
Dr. Utkarsh Seetha, Data Infosys Limited, India
Mrs. Priti Maheshwary, Maulana Azad National Institute of Technology, Bhopal
Dr. (Mrs) Padmavathi Ganapathi, Avinashilingam University for Women, Coimbatore
Assist. Prof. A. Neela madheswari, Anna university, India

Prof. Ganesan Ramachandra Rao, PSG College of Arts and Science, India
Mr. Kamanashis Biswas, Daffodil International University, Bangladesh
Dr. Atul Gonsai, Saurashtra University, Gujarat, India
Mr. Angkoon Phinyomark, Prince of Songkla University, Thailand
Mrs. G. Nalini Priya, Anna University, Chennai
Dr. P. Subashini, Avinashilingam University for Women, India
Assoc. Prof. Vijay Kumar Chakka, Dhirubhai Ambani IICT, Gandhinagar, Gujarat
Mr Jitendra Agrawal, : Rajiv Gandhi Proudhyogiki Vishwavidyalaya, Bhopal
Mr. Vishal Goyal, Department of Computer Science, Punjabi University, India
Dr. R. Baskaran, Department of Computer Science and Engineering, Anna University, Chennai
Assist. Prof, Kanwalvir Singh Dhindsa, B.B.S.B.Engg.College, Fatehgarh Sahib (Punjab), India
Dr. Jamal Ahmad Dargham, School of Engineering and Information Technology, Universiti Malaysia Sabah
Mr. Nitin Bhatia, DAV College, India
Dr. Dhavachelvan Ponnurangam, Pondicherry Central University, India
Dr. Mohd Faizal Abdollah, University of Technical Malaysia, Malaysia
Assist. Prof. Sonal Chawla, Panjab University, India
Dr. Abdul Wahid, AKG Engg. College, Ghaziabad, India
Mr. Arash Habibi Lashkari, University of Malaya (UM), Malaysia
Mr. Md. Rajibul Islam, Ibnu Sina Institute, University Technology Malaysia
Professor Dr. Sabu M. Thampi, .B.S Institute of Technology for Women, Kerala University, India
Mr. Noor Muhammed Nayeem, Université Lumière Lyon 2, 69007 Lyon, France
Dr. Himanshu Aggarwal, Department of Computer Engineering, Punjabi University, India
Prof R. Naidoo, Dept of Mathematics/Center for Advanced Computer Modelling, Durban University of
Technology, Durban, South Africa
Prof. Mydhili K Nair, Visweswaraiah Technological University, Bangalore, India
M. Prabu, Adhiyamaan College of Engineering/Anna University, India
Mr. Swakkhar Shatabda, United International University, Bangladesh
Dr. Abdur Rashid Khan, ICIT, Gomal University, Dera Ismail Khan, Pakistan
Mr. H. Abdul Shabeer, I-Nautix Technologies, Chennai, India
Dr. M. Aramudhan, Perunthalaivar Kamarajar Institute of Engineering and Technology, India
Dr. M. P. Thapliyal, Department of Computer Science, HNB Garhwal University (Central University), India
Dr. Shahaboddin Shamshirband, Islamic Azad University, Iran
Mr. Zeashan Hameed Khan, Université de Grenoble, France
Prof. Anil K Ahlawat, Ajay Kumar Garg Engineering College, Ghaziabad, UP Technical University, Lucknow
Mr. Longe Olumide Babatope, University Of Ibadan, Nigeria
Associate Prof. Raman Maini, University College of Engineering, Punjabi University, India
Dr. Maslin Masrom, University Technology Malaysia, Malaysia
Sudipta Chattopadhyay, Jadavpur University, Kolkata, India
Dr. Dang Tuan NGUYEN, University of Information Technology, Vietnam National University - Ho Chi Minh
City
Dr. Mary Lourde R., BITS-PILANI Dubai, UAE
Dr. Abdul Aziz, University of Central Punjab, Pakistan

Mr. Karan Singh, Gautam Budtha University, India
Mr. Avinash Pokhriyal, Uttar Pradesh Technical University, Lucknow, India
Associate Prof Dr Zuraini Ismail, University Technology Malaysia, Malaysia
Assistant Prof. Yasser M. Alginahi, Taibah University, Madinah Munawwarah, KSA
Mr. Dakshina Ranjan Kisku, West Bengal University of Technology, India
Mr. Raman Kumar, Dr B R Ambedkar National Institute of Technology, Jalandhar, Punjab, India
Associate Prof. Samir B. Patel, Institute of Technology, Nirma University, India
Dr. M.Munir Ahamed Rabbani, B. S. Abdur Rahman University, India
Asst. Prof. Koushik Majumder, West Bengal University of Technology, India
Dr. Alex Pappachen James, Queensland Micro-nanotechnology center, Griffith University, Australia
Assistant Prof. S. Hariharan, B.S. Abdur Rahman University, India
Asst Prof. Jasmine. K. S, R.V.College of Engineering, India
Mr Naushad Ali Mamode Khan, Ministry of Education and Human Resources, Mauritius
Prof. Mahesh Goyani, G H Patel Collge of Engg. & Tech, V.V.N, Anand, Gujarat, India
Dr. Mana Mohammed, University of Tlemcen, Algeria
Prof. Jatinder Singh, Universal Institutiion of Engg. & Tech. CHD, India
Mrs. M. Anandhavalli Gauthaman, Sikkim Manipal Institute of Technology, Majitar, East Sikkim
Dr. Bin Guo, Institute Telecom SudParis, France
Mrs. Maleika Mehr Nigar Mohamed Heenaye-Mamode Khan, University of Mauritius
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Mr. V. Bala Dhandayuthapani, Mekelle University, Ethiopia
Dr. Irfan Syamsuddin, State Polytechnic of Ujung Pandang, Indonesia
Mr. Kavi Kumar Khedo, University of Mauritius, Mauritius
Mr. Ravi Chandiran, Zagro Singapore Pte Ltd. Singapore
Mr. Milindkumar V. Sarode, Jawaharlal Darda Institute of Engineering and Technology, India
Dr. Shamimul Qamar, KSJ Institute of Engineering & Technology, India
Dr. C. Arun, Anna University, India
Assist. Prof. M.N.Birje, Basaveshwar Engineering College, India
Prof. Hamid Reza Naji, Department of Computer Enigneering, Shahid Beheshti University, Tehran, Iran
Assist. Prof. Debasis Giri, Department of Computer Science and Engineering, Haldia Institute of Technology
Subhabrata Barman, Haldia Institute of Technology, West Bengal
Mr. M. I. Lali, COMSATS Institute of Information Technology, Islamabad, Pakistan
Dr. Feroz Khan, Central Institute of Medicinal and Aromatic Plants, Lucknow, India
Mr. R. Nagendran, Institute of Technology, Coimbatore, Tamilnadu, India
Mr. Amnach Khawne, King Mongkut's Institute of Technology Ladkrabang, Ladkrabang, Bangkok, Thailand
Dr. P. Chakrabarti, Sir Padampat Singhanian University, Udaipur, India
Mr. Nafiz Imtiaz Bin Hamid, Islamic University of Technology (IUT), Bangladesh.
Shahab-A. Shamshirband, Islamic Azad University, Chalous, Iran
Prof. B. Priestly Shan, Anna Univeristy, Tamilnadu, India
Venkatramreddy Velma, Dept. of Bioinformatics, University of Mississippi Medical Center, Jackson MS USA
Akshi Kumar, Dept. of Computer Engineering, Delhi Technological University, India
Dr. Umesh Kumar Singh, Vikram University, Ujjain, India

Mr. Serguei A. Mokhov, Concordia University, Canada
Mr. Lai Khin Wee, Universiti Teknologi Malaysia, Malaysia
Dr. Awadhesh Kumar Sharma, Madan Mohan Malviya Engineering College, India
Mr. Syed R. Rizvi, Analytical Services & Materials, Inc., USA
Dr. S. Karthik, SNS College of Technology, India
Mr. Syed Qasim Bukhari, CIMET (Universidad de Granada), Spain
Mr. A.D.Potgantwar, Pune University, India
Dr. Himanshu Aggarwal, Punjabi University, India
Mr. Rajesh Ramachandran, Naipunya Institute of Management and Information Technology, India
Dr. K.L. Shunmuganathan, R.M.K Engg College, Kavaraipeetai, Chennai
Dr. Prasant Kumar Pattnaik, KIST, India.
Dr. Ch. Aswani Kumar, VIT University, India
Mr. Ijaz Ali Shoukat, King Saud University, Riyadh KSA
Mr. Arun Kumar, Sir Padam Pat Singhania University, Udaipur, Rajasthan
Mr. Muhammad Imran Khan, Universiti Teknologi PETRONAS, Malaysia
Dr. Natarajan Meghanathan, Jackson State University, Jackson, MS, USA
Mr. Mohd Zaki Bin Mas'ud, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia
Prof. Dr. R. Geetharamani, Dept. of Computer Science and Eng., Rajalakshmi Engineering College, India
Dr. Smita Rajpal, Institute of Technology and Management, Gurgaon, India
Dr. S. Abdul Khader Jilani, University of Tabuk, Tabuk, Saudi Arabia
Mr. Syed Jamal Haider Zaidi, Bahria University, Pakistan
Dr. N. Devarajan, Government College of Technology, Coimbatore, Tamilnadu, INDIA
Mr. R. Jagadeesh Kannan, RMK Engineering College, India
Mr. Deo Prakash, Shri Mata Vaishno Devi University, India
Mr. Mohammad Abu Naser, Dept. of EEE, IUT, Gazipur, Bangladesh
Assist. Prof. Prasun Ghosal, Bengal Engineering and Science University, India
Mr. Md. Golam Kaosar, School of Engineering and Science, Victoria University, Melbourne City, Australia
Mr. R. Mahammad Shafi, Madanapalle Institute of Technology & Science, India
Dr. F.Sagayaraj Francis, Pondicherry Engineering College, India
Dr. Ajay Goel, HIET, Kaithal, India
Mr. Nayak Sunil Kashibarao, Bahirji Smarak Mahavidyalaya, India
Mr. Suhas J Manangi, Microsoft India
Dr. Kalyankar N. V., Yeshwant Mahavidyalaya, Nanded, India
Dr. K.D. Verma, S.V. College of Post graduate studies & Research, India
Dr. Amjad Rehman, University Technology Malaysia, Malaysia
Mr. Rachit Garg, L K College, Jalandhar, Punjab
Mr. J. William, M.A.M college of Engineering, Trichy, Tamilnadu, India
Prof. Jue-Sam Chou, Nanhua University, College of Science and Technology, Taiwan
Dr. Thorat S.B., Institute of Technology and Management, India
Mr. Ajay Prasad, Sir Padampat Singhania University, Udaipur, India
Dr. Kamaljit I. Lakhtaria, Atmiya Institute of Technology & Science, India
Mr. Syed Rafiul Hussain, Ahsanullah University of Science and Technology, Bangladesh

Mrs Fazeela Tunnisa, Najran University, Kingdom of Saudi Arabia
Mrs Kavita Taneja, Maharishi Markandeshwar University, Haryana, India
Mr. Maniyar Shiraz Ahmed, Najran University, Najran, KSA
Mr. Anand Kumar, AMC Engineering College, Bangalore
Dr. Rakesh Chandra Gangwar, Beant College of Engg. & Tech., Gurdaspur (Punjab) India
Dr. V V Rama Prasad, Sree Vidyanikethan Engineering College, India
Assist. Prof. Neetesh Kumar Gupta, Technocrats Institute of Technology, Bhopal (M.P.), India
Mr. Ashish Seth, Uttar Pradesh Technical University, Lucknow, UP India
Dr. V V S S S Balaram, Sreenidhi Institute of Science and Technology, India
Mr Rahul Bhatia, Lingaya's Institute of Management and Technology, India
Prof. Niranjana Reddy, P, KITS, Warangal, India
Prof. Rakesh. Lingappa, Vijetha Institute of Technology, Bangalore, India
Dr. Mohammed Ali Hussain, Nimra College of Engineering & Technology, Vijayawada, A.P., India
Dr. A.Srinivasan, MNM Jain Engineering College, Rajiv Gandhi Salai, Thorapakkam, Chennai
Mr. Rakesh Kumar, M.M. University, Mullana, Ambala, India
Dr. Lena Khaled, Zarqa Private University, Aman, Jordan
Ms. Supriya Kapoor, Patni/Lingaya's Institute of Management and Tech., India
Dr. Tossapon Boongoen, Aberystwyth University, UK
Dr. Bilal Alatas, Firat University, Turkey
Assist. Prof. Jyoti Praakash Singh, Academy of Technology, India
Dr. Ritu Soni, GNG College, India
Dr. Mahendra Kumar, Sagar Institute of Research & Technology, Bhopal, India.
Dr. Binod Kumar, Lakshmi Narayan College of Tech.(LNCT)Bhopal India
Dr. Muzhir Shaban Al-Ani, Amman Arab University Amman – Jordan
Dr. T.C. Manjunath, ATRIA Institute of Tech, India
Mr. Muhammad Zakarya, COMSATS Institute of Information Technology (CIIT), Pakistan
Assist. Prof. Harmunish Taneja, M. M. University, India
Dr. Chitra Dhawale, SICSR, Model Colony, Pune, India
Mrs Sankari Muthukaruppan, Nehru Institute of Engineering and Technology, Anna University, India
Mr. Aaqif Afzaal Abbasi, National University Of Sciences And Technology, Islamabad
Prof. Ashutosh Kumar Dubey, Trinity Institute of Technology and Research Bhopal, India
Mr. G. Appasami, Dr. Pauls Engineering College, India
Mr. M Yasin, National University of Science and Tech, Karachi (NUST), Pakistan
Mr. Yaser Miaji, University Utara Malaysia, Malaysia
Mr. Shah Ahsanul Haque, International Islamic University Chittagong (IIUC), Bangladesh
Prof. (Dr) Syed Abdul Sattar, Royal Institute of Technology & Science, India
Dr. S. Sasikumar, Roever Engineering College
Assist. Prof. Monit Kapoor, Maharishi Markandeshwar University, India
Mr. Nwaocha Vivian O, National Open University of Nigeria
Dr. M. S. Vijaya, GR Govindarajulu School of Applied Computer Technology, India
Assist. Prof. Chakresh Kumar, Manav Rachna International University, India
Mr. Kunal Chadha, R&D Software Engineer, Gemalto, Singapore

Mr. Mueen Uddin, Universiti Teknologi Malaysia, UTM , Malaysia
Dr. Dhuha Basheer abdullah, Mosul university, Iraq
Mr. S. Audithan, Annamalai University, India
Prof. Vijay K Chaudhari, Technocrats Institute of Technology , India
Associate Prof. Mohd Ilyas Khan, Technocrats Institute of Technology , India
Dr. Vu Thanh Nguyen, University of Information Technology, HoChiMinh City, VietNam
Assist. Prof. Anand Sharma, MITS, Lakshmangarh, Sikar, Rajasthan, India
Prof. T V Narayana Rao, HITAM Engineering college, Hyderabad
Mr. Deepak Gour, Sir Padampat Singhania University, India
Assist. Prof. Amutharaj Joyson, Kalasalingam University, India
Mr. Ali Balador, Islamic Azad University, Iran
Mr. Mohit Jain, Maharaja Surajmal Institute of Technology, India
Mr. Dilip Kumar Sharma, GLA Institute of Technology & Management, India
Dr. Debojyoti Mitra, Sir padampat Singhania University, India
Dr. Ali Dehghantanha, Asia-Pacific University College of Technology and Innovation, Malaysia
Mr. Zhao Zhang, City University of Hong Kong, China
Prof. S.P. Setty, A.U. College of Engineering, India
Prof. Patel Rakeshkumar Kantilal, Sankalchand Patel College of Engineering, India
Mr. Biswajit Bhowmik, Bengal College of Engineering & Technology, India
Mr. Manoj Gupta, Apex Institute of Engineering & Technology, India
Assist. Prof. Ajay Sharma, Raj Kumar Goel Institute Of Technology, India
Assist. Prof. Ramveer Singh, Raj Kumar Goel Institute of Technology, India
Dr. Hanan Elazhary, Electronics Research Institute, Egypt
Dr. Hosam I. Faiq, USM, Malaysia
Prof. Dipti D. Patil, MAEER's MIT College of Engg. & Tech, Pune, India
Assist. Prof. Devendra Chack, BCT Kumaon engineering College Dwarahat Almora, India
Prof. Manpreet Singh, M. M. Engg. College, M. M. University, India
Assist. Prof. M. Sadiq ali Khan, University of Karachi, Pakistan
Mr. Prasad S. Halgaonkar, MIT - College of Engineering, Pune, India
Dr. Imran Ghani, Universiti Teknologi Malaysia, Malaysia
Prof. Varun Kumar Kakar, Kumaon Engineering College, Dwarahat, India
Assist. Prof. Nisheeth Joshi, Apaji Institute, Banasthali University, Rajasthan, India
Associate Prof. Kunwar S. Vaisla, VCT Kumaon Engineering College, India
Prof Anupam Choudhary, Bhilai School Of Engg.,Bhilai (C.G.),India
Mr. Divya Prakash Shrivastava, Al Jabal Al garbi University, Zawya, Libya
Associate Prof. Dr. V. Radha, Avinashilingam Deemed university for women, Coimbatore.
Dr. Kasarapu Ramani, JNT University, Anantapur, India
Dr. Anuraag Awasthi, Jayoti Vidyapeeth Womens University, India
Dr. C G Ravichandran, R V S College of Engineering and Technology, India
Dr. Mohamed A. Deriche, King Fahd University of Petroleum and Minerals, Saudi Arabia
Mr. Abbas Karimi, Universiti Putra Malaysia, Malaysia
Mr. Amit Kumar, Jaypee University of Engg. and Tech., India

Dr. Nikolai Stoianov, Defense Institute, Bulgaria
Assist. Prof. S. Ranichandra, KSR College of Arts and Science, Tiruchencode
Mr. T.K.P. Rajagopal, Diamond Horse International Pvt Ltd, India
Dr. Md. Ekramul Hamid, Rajshahi University, Bangladesh
Mr. Hemanta Kumar Kalita , TATA Consultancy Services (TCS), India
Dr. Messaouda Azzouzi, Ziane Achour University of Djelfa, Algeria
Prof. (Dr.) Juan Jose Martinez Castillo, "Gran Mariscal de Ayacucho" University and Acantelys research Group, Venezuela
Dr. Jatinderkumar R. Saini, Narmada College of Computer Application, India
Dr. Babak Bashari Rad, University Technology of Malaysia, Malaysia
Dr. Nighat Mir, Effat University, Saudi Arabia
Prof. (Dr.) G.M.Nasira, Sasurie College of Engineering, India
Mr. Varun Mittal, Gemalto Pte Ltd, Singapore
Assist. Prof. Mrs P. Banumathi, Kathir College Of Engineering, Coimbatore
Assist. Prof. Quan Yuan, University of Wisconsin-Stevens Point, US
Dr. Pranam Paul, Narula Institute of Technology, Agarpara, West Bengal, India
Assist. Prof. J. Ramkumar, V.L.B Janakiammal college of Arts & Science, India
Mr. P. Sivakumar, Anna university, Chennai, India
Mr. Md. Humayun Kabir Biswas, King Khalid University, Kingdom of Saudi Arabia
Mr. Mayank Singh, J.P. Institute of Engg & Technology, Meerut, India
HJ. Kamaruzaman Jusoff, Universiti Putra Malaysia
Mr. Nikhil Patrick Lobo, CADES, India
Dr. Amit Wason, Rayat-Bahra Institute of Engineering & Boi-Technology, India
Dr. Rajesh Shrivastava, Govt. Benazir Science & Commerce College, Bhopal, India
Assist. Prof. Vishal Bharti, DCE, Gurgaon
Mrs. Sunita Bansal, Birla Institute of Technology & Science, India
Dr. R. Sudhakar, Dr.Mahalingam college of Engineering and Technology, India
Dr. Amit Kumar Garg, Shri Mata Vaishno Devi University, Katra(J&K), India
Assist. Prof. Raj Gaurang Tiwari, AZAD Institute of Engineering and Technology, India
Mr. Hamed Taherdoost, Tehran, Iran
Mr. Amin Daneshmand Malayeri, YRC, IAU, Malayer Branch, Iran
Mr. Shantanu Pal, University of Calcutta, India
Dr. Terry H. Walcott, E-Promag Consultancy Group, United Kingdom
Dr. Ezekiel U OKIKE, University of Ibadan, Nigeria
Mr. P. Mahalingam, Caledonian College of Engineering, Oman
Dr. Mahmoud M. A. Abd Ellatif, Mansoura University, Egypt
Prof. Kunwar S. Vaisla, BCT Kumaon Engineering College, India
Prof. Mahesh H. Panchal, Kalol Institute of Technology & Research Centre, India
Mr. Muhammad Asad, Technical University of Munich, Germany
Mr. AliReza Shams Shafigh, Azad Islamic university, Iran
Prof. S. V. Nagaraj, RMK Engineering College, India
Mr. Ashikali M Hasan, Senior Researcher, CeINet security, India

Dr. Adnan Shahid Khan, University Technology Malaysia, Malaysia
Mr. Prakash Gajanan Burade, Nagpur University/ITM college of engg, Nagpur, India
Dr. Jagdish B.Helonde, Nagpur University/ITM college of engg, Nagpur, India
Professor, Doctor BOUHORMA Mohammed, Univertsity Abdelmalek Essaadi, Morocco
Mr. K. Thirumalaivasan, Pondicherry Engg. College, India
Mr. Umbarkar Anantkumar Janardan, Walchand College of Engineering, India
Mr. Ashish Chaurasia, Gyan Ganga Institute of Technology & Sciences, India
Mr. Sunil Taneja, Kurukshetra University, India
Mr. Fauzi Adi Rafrastara, Dian Nuswantoro University, Indonesia
Dr. Yaduvir Singh, Thapar University, India
Dr. Ioannis V. Koskosas, University of Western Macedonia, Greece
Dr. Vasantha Kalyani David, Avinashilingam University for women, Coimbatore
Dr. Ahmed Mansour Manasrah, Universiti Sains Malaysia, Malaysia
Miss. Nazanin Sadat Kazazi, University Technology Malaysia, Malaysia
Mr. Saeed Rasouli Heikalabad, Islamic Azad University - Tabriz Branch, Iran
Assoc. Prof. Dharendra Mishra, SVKM's NMIMS University, India
Prof. Shapoor Zarei, UAE Inventors Association, UAE
Prof. B.Raja Sarath Kumar, Lenora College of Engineering, India
Dr. Bashir Alam, Jamia millia Islamia, Delhi, India
Prof. Anant J Umbarkar, Walchand College of Engg., India
Assist. Prof. B. Bharathi, Sathyabama University, India
Dr. Fokrul Alom Mazarbhuiya, King Khalid University, Saudi Arabia
Prof. T.S.Jeyali Laseeth, Anna University of Technology, Tirunelveli, India
Dr. M. Balraju, Jawahar Lal Nehru Technological University Hyderabad, India
Dr. Vijayalakshmi M. N., R.V.College of Engineering, Bangalore
Prof. Walid Moudani, Lebanese University, Lebanon
Dr. Saurabh Pal, VBS Purvanchal University, Jaunpur, India
Associate Prof. Suneet Chaudhary, Dehradun Institute of Technology, India
Associate Prof. Dr. Manuj Darbari, BBD University, India
Ms. Prema Selvaraj, K.S.R College of Arts and Science, India
Assist. Prof. Ms.S.Sasikala, KSR College of Arts & Science, India
Mr. Sukhvinder Singh Deora, NC Institute of Computer Sciences, India
Dr. Abhay Bansal, Amity School of Engineering & Technology, India
Ms. Sumita Mishra, Amity School of Engineering and Technology, India
Professor S. Viswanadha Raju, JNT University Hyderabad, India
Mr. Asghar Shahrzad Khashandarag, Islamic Azad University Tabriz Branch, India
Mr. Manoj Sharma, Panipat Institute of Engg. & Technology, India
Mr. Shakeel Ahmed, King Faisal University, Saudi Arabia
Dr. Mohamed Ali Mahjoub, Institute of Engineer of Monastir, Tunisia
Mr. Adri Jovin J.J., SriGuru Institute of Technology, India
Dr. Sukumar Senthilkumar, Universiti Sains Malaysia, Malaysia
Mr. Rakesh Bharati, Dehradun Institute of Technology Dehradun, India

Mr. Shervan Fekri Ershad, Shiraz International University, Iran
Mr. Md. Safiqul Islam, Daffodil International University, Bangladesh
Mr. Mahmudul Hasan, Daffodil International University, Bangladesh
Prof. Mandakini Tayade, UIT, RGTU, Bhopal, India
Ms. Sarla More, UIT, RGTU, Bhopal, India
Mr. Tushar Hrishikesh Jaware, R.C. Patel Institute of Technology, Shirpur, India
Ms. C. Divya, Dr G R Damodaran College of Science, Coimbatore, India
Mr. Fahimuddin Shaik, Annamacharya Institute of Technology & Sciences, India
Dr. M. N. Giri Prasad, JNTUCE, Pulivendula, A.P., India
Assist. Prof. Chintan M Bhatt, Charotar University of Science And Technology, India
Prof. Sahista Machchhar, Marwadi Education Foundation's Group of institutions, India
Assist. Prof. Navnish Goel, S. D. College Of Enginnering & Technology, India
Mr. Khaja Kamaluddin, Sirt University, Sirt, Libya
Mr. Mohammad Zaidul Karim, Daffodil International, Bangladesh
Mr. M. Vijayakumar, KSR College of Engineering, Tiruchengode, India
Mr. S. A. Ahsan Rajon, Khulna University, Bangladesh
Dr. Muhammad Mohsin Nazir, LCW University Lahore, Pakistan
Mr. Mohammad Asadul Hoque, University of Alabama, USA
Mr. P.V.Sarathchand, Indur Institute of Engineering and Technology, India
Mr. Durgesh Samadhiya, Chung Hua University, Taiwan
Dr Venu Kuthadi, University of Johannesburg, Johannesburg, RSA
Dr. (Er) Jasvir Singh, Guru Nanak Dev University, Amritsar, Punjab, India
Mr. Jasmin Cosic, Min. of the Interior of Una-sana canton, B&H, Bosnia and Herzegovina
Dr S. Rajalakshmi, Botho College, South Africa
Dr. Mohamed Sarrab, De Montfort University, UK
Mr. Basappa B. Kodada, Canara Engineering College, India
Assist. Prof. K. Ramana, Annamacharya Institute of Technology and Sciences, India
Dr. Ashu Gupta, Apeejay Institute of Management, Jalandhar, India
Assist. Prof. Shaik Rasool, Shadan College of Engineering & Technology, India
Assist. Prof. K. Suresh, Annamacharya Institute of Tech & Sci. Rajampet, AP, India
Dr . G. Singaravel, K.S.R. College of Engineering, India
Dr B. G. Geetha, K.S.R. College of Engineering, India
Assist. Prof. Kavita Choudhary, ITM University, Gurgaon
Dr. Mehrdad Jalali, Azad University, Mashhad, Iran
Megha Goel, Shamli Institute of Engineering and Technology, Shamli, India
Mr. Chi-Hua Chen, Institute of Information Management, National Chiao-Tung University, Taiwan (R.O.C.)
Assoc. Prof. A. Rajendran, RVS College of Engineering and Technology, India
Assist. Prof. S. Jaganathan, RVS College of Engineering and Technology, India
Assoc. Prof. (Dr.) A S N Chakravarthy, JNTUK University College of Engineering Vizianagaram (State University)
Assist. Prof. Deepshikha Patel, Technocrat Institute of Technology, India
Assist. Prof. Maram Balajee, GMRIT, India

Assist. Prof. Monika Bhatnagar, TIT, India
Prof. Gaurang Panchal, Charotar University of Science & Technology, India
Prof. Anand K. Tripathi, Computer Society of India
Prof. Jyoti Chaudhary, High Performance Computing Research Lab, India
Assist. Prof. Supriya Raheja, ITM University, India
Dr. Pankaj Gupta, Microsoft Corporation, U.S.A.
Assist. Prof. Panchamukesh Chandaka, Hyderabad Institute of Tech. & Management, India
Prof. Mohan H.S, SJB Institute Of Technology, India
Mr. Hossein Malekinezhad, Islamic Azad University, Iran
Mr. Zatin Gupta, Universti Malaysia, Malaysia
Assist. Prof. Amit Chauhan, Phonics Group of Institutions, India
Assist. Prof. Ajal A. J., METS School Of Engineering, India
Mrs. Omowunmi Omobola Adeyemo, University of Ibadan, Nigeria
Dr. Bharat Bhushan Agarwal, I.F.T.M. University, India
Md. Nazrul Islam, University of Western Ontario, Canada
Tushar Kanti, L.N.C.T, Bhopal, India
Er. Aumreesh Kumar Saxena, SIRTs College Bhopal, India
Mr. Mohammad Monirul Islam, Daffodil International University, Bangladesh
Dr. Kashif Nisar, University Utara Malaysia, Malaysia
Dr. Wei Zheng, Rutgers Univ/ A10 Networks, USA
Associate Prof. Rituraj Jain, Vyas Institute of Engg & Tech, Jodhpur – Rajasthan
Assist. Prof. Apoorvi Sood, I.T.M. University, India
Dr. Kayhan Zrar Ghafoor, University Technology Malaysia, Malaysia
Mr. Swapnil Soner, Truba Institute College of Engineering & Technology, Indore, India
Ms. Yogita Gigras, I.T.M. University, India
Associate Prof. Neelima Sadineni, Pydha Engineering College, India Pydha Engineering College
Assist. Prof. K. Deepika Rani, HITAM, Hyderabad
Ms. Shikha Maheshwari, Jaipur Engineering College & Research Centre, India
Prof. Dr V S Giridhar Akula, Avanthi's Scientific Tech. & Research Academy, Hyderabad
Prof. Dr.S.Saravanan, Muthayammal Engineering College, India
Mr. Mehdi Golsorkhatabar Amiri, Islamic Azad University, Iran
Prof. Amit Sadanand Savyanavar, MITCOE, Pune, India
Assist. Prof. P.Oliver Jayaprakash, Anna University, Chennai
Assist. Prof. Ms. Sujata, ITM University, Gurgaon, India
Dr. Asoke Nath, St. Xavier's College, India
Mr. Masoud Rafighi, Islamic Azad University, Iran
Assist. Prof. RamBabu Pemula, NIMRA College of Engineering & Technology, India
Assist. Prof. Ms Rita Chhikara, ITM University, Gurgaon, India
Mr. Sandeep Maan, Government Post Graduate College, India
Prof. Dr. S. Muralidharan, Mepco Schlenk Engineering College, India
Associate Prof. T.V.Sai Krishna, QIS College of Engineering and Technology, India
Mr. R. Balu, Bharathiar University, Coimbatore, India

Assist. Prof. Shekhar. R, Dr.SM College of Engineering, India
Prof. P. Senthikumar, Vivekanandha Institute of Engineering and Technology for Woman, India
Mr. M. Kamarajan, PSNA College of Engineering & Technology, India
Dr. Angajala Srinivasa Rao, Jawaharlal Nehru Technical University, India
Assist. Prof. C. Venkatesh, A.I.T.S, Rajampet, India
Mr. Afshin Rezakhani Roozbahani, Ayatollah Boroujerdi University, Iran
Mr. Laxmi chand, SCTL, Noida, India
Dr. Dr. Abdul Hannan, Vivekanand College, Aurangabad
Prof. Mahesh Panchal, KITRC, Gujarat
Dr. A. Subramani, K.S.R. College of Engineering, Tiruchengode
Assist. Prof. Prakash M, Rajalakshmi Engineering College, Chennai, India
Assist. Prof. Akhilesh K Sharma, Sir Padampat Singhania University, India
Ms. Varsha Sahni, Guru Nanak Dev Engineering College, Ludhiana, India
Associate Prof. Trilochan Rout, NM Institute of Engineering and Technology, India
Mr. Srikanta Kumar Mohapatra, NMIET, Orissa, India
Mr. Waqas Haider Bangyal, Iqra University Islamabad, Pakistan
Dr. S. Vijayaragavan, Christ College of Engineering and Technology, Pondicherry, India
Prof. Elboukhari Mohamed, University Mohammed First, Oujda, Morocco
Dr. Muhammad Asif Khan, King Faisal University, Saudi Arabia
Dr. Nagy Ramadan Darwish Omran, Cairo University, Egypt.
Assistant Prof. Anand Nayyar, KCL Institute of Management and Technology, India
Mr. G. Premsankar, Ericsson, India
Assist. Prof. T. Hemalatha, VELS University, India
Prof. Tejaswini Apte, University of Pune, India
Dr. Edmund Ng Giap Weng, Universiti Malaysia Sarawak, Malaysia
Mr. Mahdi Nouri, Iran University of Science and Technology, Iran
Associate Prof. S. Asif Hussain, Annamacharya Institute of technology & Sciences, India
Mrs. Kavita Pabreja, Maharaja Surajmal Institute (an affiliate of GGSIP University), India
Mr. Vorugunti Chandra Sekhar, DA-IICT, India
Mr. Muhammad Najmi Ahmad Zabidi, Universiti Teknologi Malaysia, Malaysia
Dr. Aderemi A. Atayero, Covenant University, Nigeria
Assist. Prof. Osama Sohaib, Balochistan University of Information Technology, Pakistan
Assist. Prof. K. Suresh, Annamacharya Institute of Technology and Sciences, India
Mr. Hassen Mohammed Abdulllah Alsafi, International Islamic University Malaysia (IIUM) Malaysia
Mr. Robail Yasrab, Virtual University of Pakistan, Pakistan
Mr. R. Balu, Bharathiar University, Coimbatore, India
Prof. Anand Nayyar, KCL Institute of Management and Technology, Jalandhar
Assoc. Prof. Vivek S Deshpande, MIT College of Engineering, India
Prof. K. Saravanan, Anna university Coimbatore, India
Dr. Ravendra Singh, MJP Rohilkhand University, Bareilly, India
Mr. V. Mathivanan, IBRA College of Technology, Sultanate of OMAN
Assoc. Prof. S. Asif Hussain, AITS, India

Assist. Prof. C. Venkatesh, AITS, India
Mr. Sami Ulhaq, SZABIST Islamabad, Pakistan
Dr. B. Justus Rabi, Institute of Science & Technology, India
Mr. Anuj Kumar Yadav, Dehradun Institute of technology, India
Mr. Alejandro Mosquera, University of Alicante, Spain
Assist. Prof. Arjun Singh, Sir Padampat Singhania University (SPSU), Udaipur, India
Dr. Smriti Agrawal, JB Institute of Engineering and Technology, Hyderabad
Assist. Prof. Swathi Sambangi, Visakha Institute of Engineering and Technology, India
Ms. Prabhjot Kaur, Guru Gobind Singh Indraprastha University, India
Mrs. Samaher AL-Hothali, Yanbu University College, Saudi Arabia
Prof. Rajneeshkaur Bedi, MIT College of Engineering, Pune, India
Mr. Hassen Mohammed Abdullallah Alsafi, International Islamic University Malaysia (IIUM)
Dr. Wei Zhang, Amazon.com, Seattle, WA, USA
Mr. B. Santhosh Kumar, C S I College of Engineering, Tamil Nadu
Dr. K. Reji Kumar, , N S S College, Pandalam, India
Assoc. Prof. K. Seshadri Sastry, EILM University, India
Mr. Kai Pan, UNC Charlotte, USA
Mr. Ruikar Sachin, SGGSIET, India
Prof. (Dr.) Vinodani Katiyar, Sri Ramswaroop Memorial University, India
Assoc. Prof., M. Giri, Sreenivasa Institute of Technology and Management Studies, India
Assoc. Prof. Labib Francis Gergis, Misr Academy for Engineering and Technology (MET), Egypt
Assist. Prof. Amanpreet Kaur, ITM University, India
Assist. Prof. Anand Singh Rajawat, Shri Vaishnav Institute of Technology & Science, Indore
Mrs. Hadeel Saleh Haj Aliwi, Universiti Sains Malaysia (USM), Malaysia
Dr. Abhay Bansal, Amity University, India
Dr. Mohammad A. Mezher, Fahad Bin Sultan University, KSA
Assist. Prof. Nidhi Arora, M.C.A. Institute, India
Prof. Dr. P. Suresh, Karpagam College of Engineering, Coimbatore, India
Dr. Kannan Balasubramanian, Mepco Schlenk Engineering College, India
Dr. S. Sankara Gomathi, Panimalar Engineering college, India
Prof. Anil kumar Suthar, Gujarat Technological University, L.C. Institute of Technology, India
Assist. Prof. R. Hubert Rajan, NOORUL ISLAM UNIVERSITY, India
Assist. Prof. Dr. Jyoti Mahajan, College of Engineering & Technology
Assist. Prof. Homam Reda El-Taj, College of Network Engineering, Saudi Arabia & Malaysia
Mr. Bijan Paul, Shahjalal University of Science & Technology, Bangladesh
Assoc. Prof. Dr. Ch V Phani Krishna, KL University, India
Dr. Vishal Bhatnagar, Ambedkar Institute of Advanced Communication Technologies & Research, India
Dr. Lamri LAOUAMER, Al Qassim University, Dept. Info. Systems & European University of Brittany, Dept. Computer Science, UBO, Brest, France
Prof. Ashish Babanrao Sasankar, G.H.Raisoni Institute Of Information Technology, India
Prof. Pawan Kumar Goel, Shamli Institute of Engineering and Technology, India
Mr. Ram Kumar Singh, S.V Subharti University, India

Assistant Prof. Sunish Kumar O S, Amalijothe College of Engineering, India
Dr Sanjay Bhargava, Banasthali University, India
Mr. Pankaj S. Kulkarni, AVEW's Shatabdi Institute of Technology, India
Mr. Roohollah Etemadi, Islamic Azad University, Iran
Mr. Oloruntoyin Sefiu Taiwo, Emmanuel Alayande College Of Education, Nigeria
Mr. Sumit Goyal, National Dairy Research Institute, India
Mr Jaswinder Singh Dilawari, Geeta Engineering College, India
Prof. Raghuraj Singh, Harcourt Butler Technological Institute, Kanpur
Dr. S.K. Mahendran, Anna University, Chennai, India
Dr. Amit Wason, Hindustan Institute of Technology & Management, Punjab
Dr. Ashu Gupta, Apeejay Institute of Management, India
Assist. Prof. D. Asir Antony Gnana Singh, M.I.E.T Engineering College, India
Mrs Mina Farmanbar, Eastern Mediterranean University, Famagusta, North Cyprus
Mr. Maram Balajee, GMR Institute of Technology, India
Mr. Moiz S. Ansari, Isra University, Hyderabad, Pakistan
Mr. Adebayo, Olawale Surajudeen, Federal University of Technology Minna, Nigeria
Mr. Jasvir Singh, University College Of Engg., India
Mr. Vivek Tiwari, MANIT, Bhopal, India
Assoc. Prof. R. Navaneethakrishnan, Bharathiyar College of Engineering and Technology, India
Mr. Somdip Dey, St. Xavier's College, Kolkata, India
Mr. Souleymane Balla-Arabé, Xi'an University of Electronic Science and Technology, China
Mr. Mahabub Alam, Rajshahi University of Engineering and Technology, Bangladesh
Mr. Sathyapraksh P., S.K.P Engineering College, India
Dr. N. Karthikeyan, SNS College of Engineering, Anna University, India
Dr. Binod Kumar, JSPM's, Jayawant Technical Campus, Pune, India
Assoc. Prof. Dinesh Goyal, Suresh Gyan Vihar University, India
Mr. Md. Abdul Ahad, K L University, India
Mr. Vikas Bajpai, The LNM IIT, India
Dr. Manish Kumar Anand, Salesforce (R & D Analytics), San Francisco, USA
Assist. Prof. Dheeraj Murari, Kumaon Engineering College, India
Assoc. Prof. Dr. A. Muthukumaravel, VELS University, Chennai
Mr. A. Siles Balasingh, St. Joseph University in Tanzania, Tanzania
Mr. Ravindra Daga Badgujar, R C Patel Institute of Technology, India
Dr. Preeti Khanna, SVKM's NMIMS, School of Business Management, India
Mr. Kumar Dayanand, Cambridge Institute of Technology, India
Dr. Syed Asif Ali, SMI University Karachi, Pakistan
Prof. Pallvi Pandit, Himachal Pradesh University, India
Mr. Ricardo Verschueren, University of Gloucestershire, UK
Assist. Prof. Mamta Juneja, University Institute of Engineering and Technology, Panjab University, India
Assoc. Prof. P. Surendra Varma, NRI Institute of Technology, JNTU Kakinada, India
Assist. Prof. Gaurav Shrivastava, RGPV / SVITS Indore, India
Dr. S. Sumathi, Anna University, India

Assist. Prof. Ankita M. Kapadia, Charotar University of Science and Technology, India
Mr. Deepak Kumar, Indian Institute of Technology (BHU), India
Dr. Dr. Rajan Gupta, GGSIP University, New Delhi, India
Assist. Prof M. Anand Kumar, Karpagam University, Coimbatore, India
Mr. Mr Arshad Mansoor, Pakistan Aeronautical Complex
Mr. Kapil Kumar Gupta, Ansal Institute of Technology and Management, India
Dr. Neeraj Tomer, SINE International Institute of Technology, Jaipur, India
Assist. Prof. Trunal J. Patel, C.G.Patel Institute of Technology, Uka Tarsadia University, Bardoli, Surat
Mr. Sivakumar, Codework solutions, India
Mr. Mohammad Sadegh Mirzaei, PGNR Company, Iran
Dr. Gerard G. Dumancas, Oklahoma Medical Research Foundation, USA
Mr. Varadala Sridhar, Varadhman College Engineering College, Affiliated To JNTU, Hyderabad
Assist. Prof. Manoj Dhawan, SVITS, Indore
Assoc. Prof. Chitreshh Banerjee, Suresh Gyan Vihar University, Jaipur, India
Dr. S. Santhi, SCSVMV University, India
Mr. Davood Mohammadi Souran, Ministry of Energy of Iran, Iran
Mr. Shamim Ahmed, Bangladesh University of Business and Technology, Bangladesh
Mr. Sandeep Reddivari, Mississippi State University, USA
Assoc. Prof. Ousmane Thiare, Gaston Berger University, Senegal
Dr. Hazra Imran, Athabasca University, Canada
Dr. Setu Kumar Chaturvedi, Technocrats Institute of Technology, Bhopal, India
Mr. Mohd Dilshad Ansari, Jaypee University of Information Technology, India
Ms. Jaspreet Kaur, Distance Education LPU, India
Dr. D. Nagarajan, Salalah College of Technology, Sultanate of Oman
Dr. K.V.N.R.Sai Krishna, S.V.R.M. College, India
Mr. Himanshu Pareek, Center for Development of Advanced Computing (CDAC), India
Mr. Khaldi Amine, Badji Mokhtar University, Algeria
Mr. Mohammad Sadegh Mirzaei, Scientific Applied University, Iran
Assist. Prof. Khyati Chaudhary, Ram-eesh Institute of Engg. & Technology, India
Mr. Sanjay Agal, Pacific College of Engineering Udaipur, India
Mr. Abdul Mateen Ansari, King Khalid University, Saudi Arabia
Dr. H.S. Behera, Veer Surendra Sai University of Technology (VSSUT), India
Dr. Shrikant Tiwari, Shri Shankaracharya Group of Institutions (SSGI), India
Prof. Ganesh B. Regulwar, Shri Shankarprasad Agnihotri College of Engg, India
Prof. Pinnamaneni Bhanu Prasad, Matrix vision GmbH, Germany
Dr. Shrikant Tiwari, Shri Shankaracharya Technical Campus (SSTC), India
Dr. Siddesh G.K., : Dayananada Sagar College of Engineering, Bangalore, India
Dr. Nadir Bouchama, CERIST Research Center, Algeria
Dr. R. Sathishkumar, Sri Venkateswara College of Engineering, India
Assistant Prof (Dr.) Mohamed Moussaoui, Abdelmalek Essaadi University, Morocco
Dr. S. Malathi, Panimalar Engineering College, Chennai, India
Dr. V. Subedha, Panimalar Institute of Technology, Chennai, India

Dr. Prashant Panse, Swami Vivekanand College of Engineering, Indore, India
Dr. Hamza Aldabbas, Al-Balqa'a Applied University, Jordan
Dr. G. Rasitha Banu, Vel's University, Chennai
Dr. V. D. Ambeth Kumar, Panimalar Engineering College, Chennai
Prof. Anuranjan Misra, Bhagwant Institute of Technology, Ghaziabad, India
Ms. U. Sinthuja, PSG college of arts & science, India
Dr. Ehsan Saradar Torshizi, Urmia University, Iran
Dr. Shamneesh Sharma, APG Shimla University, Shimla (H.P.), India
Assistant Prof. A. S. Syed Navaz, Muthayammal College of Arts & Science, India
Assistant Prof. Ranjit Panigrahi, Sikkim Manipal Institute of Technology, Majitar, Sikkim
Dr. Khaled Eskaf, Arab Academy for Science ,Technology & Maritime Transportation, Egypt
Dr. Nishant Gupta, University of Jammu, India
Assistant Prof. Nagarajan Sankaran, Annamalai University, Chidambaram, Tamilnadu, India
Assistant Prof. Tribikram Pradhan, Manipal Institute of Technology, India
Dr. Nasser Lotfi, Eastern Mediterranean University, Northern Cyprus
Dr. R. Manavalan, K S Rangasamy college of Arts and Science, Tamilnadu, India
Assistant Prof. P. Krishna Sankar, K S Rangasamy college of Arts and Science, Tamilnadu, India
Dr. Rahul Malik, Cisco Systems, USA
Dr. S. C. Lingareddy, ALPHA College of Engineering, India
Assistant Prof. Mohammed Shuaib, Interat University, Lucknow, India
Dr. Sachin Yele, Sanghvi Institute of Management & Science, India
Dr. T. Thambidurai, Sun Univercell, Singapore
Prof. Anandkumar Telang, BKIT, India
Assistant Prof. R. Poorvadevi, SCSVMV University, India
Dr Uttam Mande, Gitam University, India
Dr. Poornima Girish Naik, Shahu Institute of Business Education and Research (SIBER), India
Prof. Md. Abu Kausar, Jaipur National University, Jaipur, India
Dr. Mohammed Zuber, AISECT University, India
Prof. Kalum Priyanath Udagepola, King Abdulaziz University, Saudi Arabia
Dr. K. R. Ananth, Velalar College of Engineering and Technology, India
Assistant Prof. Sanjay Sharma, Roorkee Engineering & Management Institute Shamli (U.P), India
Assistant Prof. Panem Charan Arur, Priyadarshini Institute of Technology, India
Dr. Ashwak Mahmood muhsen alabaichi, Karbala University / College of Science, Iraq
Dr. Urmila Shrawankar, G H Raisonni College of Engineering, Nagpur (MS), India
Dr. Krishan Kumar Paliwal, Panipat Institute of Engineering & Technology, India
Dr. Mukesh Negi, Tech Mahindra, India
Dr. Anuj Kumar Singh, Amity University Gurgaon, India
Dr. Babar Shah, Gyeongsang National University, South Korea
Assistant Prof. Jayprakash Upadhyay, SRI-TECH Jabalpur, India
Assistant Prof. Varadala Sridhar, Vidya Jyothi Institute of Technology, India
Assistant Prof. Parameshachari B D, KSIT, Bangalore, India
Assistant Prof. Ankit Garg, Amity University, Haryana, India

Assistant Prof. Rajashe Karappa, SDM CET, Karnataka, India
Assistant Prof. Varun Jasuja, GNIT, India
Assistant Prof. Sonal Honale, Abha Gaikwad Patil College of Engineering Nagpur, India
Dr. Pooja Choudhary, CT Group of Institutions, NIT Jalandhar, India
Dr. Faouzi Hidoussi, UHL Batna, Algeria
Dr. Naseer Ali Husieen, Wasit University, Iraq
Assistant Prof. Vinod Kumar Shukla, Amity University, Dubai
Dr. Ahmed Farouk Metwaly, K L University
Mr. Mohammed Noaman Murad, Cihan University, Iraq
Dr. Suxing Liu, Arkansas State University, USA
Dr. M. Gomathi, Velalar College of Engineering and Technology, India
Assistant Prof. Sumardiono, College PGRI Blitar, Indonesia
Dr. Latika Kharb, Jagan Institute of Management Studies (JIMS), Delhi, India
Associate Prof. S. Raja, Pauls College of Engineering and Technology, Tamilnadu, India
Assistant Prof. Seyed Reza Pakize, Shahid Sani High School, Iran
Dr. Thiyagu Nagaraj, University-INOUE, India
Assistant Prof. Noreen Sarai, Harare Institute of Technology, Zimbabwe
Assistant Prof. Gajanand Sharma, Suresh Gyan Vihar University Jaipur, Rajasthan, India
Assistant Prof. Mapari Vikas Prakash, Siddhant COE, Sudumbare, Pune, India
Dr. Devesh Katiyar, Shri Ramswaroop Memorial University, India
Dr. Shenshen Liang, University of California, Santa Cruz, US
Assistant Prof. Mohammad Abu Omar, Limkokwing University of Creative Technology- Malaysia
Mr. Snehasis Banerjee, Tata Consultancy Services, India
Assistant Prof. Kibona Lusekelo, Ruaha Catholic University (RUCU), Tanzania
Assistant Prof. Adib Kabir Chowdhury, University College Technology Sarawak, Malaysia
Dr. Ying Yang, Computer Science Department, Yale University, USA
Dr. Vinay Shukla, Institute Of Technology & Management, India
Dr. Liviu Octavian Mafteiu-Scail, West University of Timisoara, Romania
Assistant Prof. Rana Khudhair Abbas Ahmed, Al-Rafidain University College, Iraq
Assistant Prof. Nitin A. Naik, S.R.T.M. University, India
Dr. Timothy Powers, University of Hertfordshire, UK
Dr. S. Prasath, Bharathiar University, Erode, India
Dr. Ritu Shrivastava, SIRTHS Bhopal, India
Prof. Rohit Shrivastava, Mittal Institute of Technology, Bhopal, India
Dr. Gianina Mihai, Dunarea de Jos" University of Galati, Romania
Assistant Prof. Ms. T. Kalai Selvi, Erode Sengunthar Engineering College, India
Assistant Prof. Ms. C. Kavitha, Erode Sengunthar Engineering College, India
Assistant Prof. K. Sinivasamoorthi, Erode Sengunthar Engineering College, India
Assistant Prof. Mallikarjun C Sarsamba Bheemna Khandre Institute Technology, Bhalki, India
Assistant Prof. Vishwanath Chikaraddi, Veermata Jijabai technological Institute (Central Technological Institute), India
Assistant Prof. Dr. Ikvinderpal Singh, Trai Shatabdi GGS Khalsa College, India

Assistant Prof. Mohammed Noaman Murad, Cihan University, Iraq
Professor Yousef Farhaoui, Moulay Ismail University, Errachidia, Morocco
Dr. Parul Verma, Amity University, India
Professor Yousef Farhaoui, Moulay Ismail University, Errachidia, Morocco
Assistant Prof. Madhavi Dhingra, Amity University, Madhya Pradesh, India
Assistant Prof.. G. Selvavinayagam, SNS College of Technology, Coimbatore, India
Assistant Prof. Madhavi Dhingra, Amity University, MP, India
Professor Kartheesan Log, Anna University, Chennai
Professor Vasudeva Acharya, Shri Madhwa vadiraja Institute of Technology, India
Dr. Asif Iqbal Hajamydeen, Management & Science University, Malaysia
Assistant Prof., Mahendra Singh Meena, Amity University Haryana
Assistant Professor Manjeet Kaur, Amity University Haryana
Dr. Mohamed Abd El-Basset Matwalli, Zagazig University, Egypt
Dr. Ramani Kannan, Universiti Teknologi PETRONAS, Malaysia
Assistant Prof. S. Jagadeesan Subramaniam, Anna University, India
Assistant Prof. Dharmendra Choudhary, Tripura University, India
Assistant Prof. Deepika Vodnala, SR Engineering College, India
Dr. Kai Cong, Intel Corporation & Computer Science Department, Portland State University, USA
Dr. Kailas R Patil, Vishwakarma Institute of Information Technology (VIIT), India
Dr. Omar A. Alzubi, Faculty of IT / Al-Balqa Applied University, Jordan
Assistant Prof. Kareemullah Shaik, Nimra Institute of Science and Technology, India
Assistant Prof. Chirag Modi, NIT Goa
Dr. R. Ramkumar, Nandha Arts And Science College, India
Dr. Priyadharshini Vydhialingam, Harathiar University, India
Dr. P. S. Jagadeesh Kumar, DBIT, Bangalore, Karnataka
Dr. Vikas Thada, AMITY University, Pachgaon
Dr. T. A. Ashok Kumar, Institute of Management, Christ University, Bangalore
Dr. Shaheera Rashwan, Informatics Research Institute
Dr. S. Preetha Gunasekar, Bharathiyar University, India
Asst Professor Sameer Dev Sharma, Uttaranchal University, Dehradun
Dr. Zhihan Iv, Chinese Academy of Science, China
Dr. Ikvinderpal Singh, Trai Shatabdi GGS Khalsa College, Amritsar
Dr. Umar Ruhi, University of Ottawa, Canada
Dr. Jasmin Cosic, University of Bihac, Bosnia and Herzegovina
Dr. Homam Reda El-Taj, University of Tabuk, Kingdom of Saudi Arabia
Dr. Mostafa Ghobaei Arani, Islamic Azad University, Iran
Dr. Ayyasamy Ayyanar, Annamalai University, India
Dr. Selvakumar Manickam, Universiti Sains Malaysia, Malaysia
Dr. Murali Krishna Namana, GITAM University, India
Dr. Smriti Agrawal, Chaitanya Bharathi Institute of Technology, Hyderabad, India
Professor Vimalathithan Rathinasabapathy, Karpagam College Of Engineering, India
Dr. Sushil Chandra Dimri, Graphic Era University, India

Dr. Dinh-Sinh Mai, Le Quy Don Technical University, Vietnam

Dr. S. Rama Sree, Aditya Engg. College, India

Dr. Ehab T. Alnaway, Sadat Academy, Egypt

Dr. Patrick D. Cerna, Haramaya University, Ethiopia

Dr. Vishal Jain, Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), India

Associate Prof. Dr. Jiliang Zhang, North Eastern University, China

Dr. Sharefa Murad, Middle East University, Jordan

Dr. Ajeet Singh Poonia, Govt. College of Engineering & technology, Rajasthan, India

Dr. Vahid Esmaeelzadeh, University of Science and Technology, Iran

CALL FOR PAPERS

International Journal of Computer Science and Information Security

IJCSIS 2016

ISSN: 1947-5500

<http://sites.google.com/site/ijcsis/>

International Journal Computer Science and Information Security, IJCSIS, is the premier scholarly venue in the areas of computer science and security issues. IJCSIS 2011 will provide a high profile, leading edge platform for researchers and engineers alike to publish state-of-the-art research in the respective fields of information technology and communication security. The journal will feature a diverse mixture of publication articles including core and applied computer science related topics.

Authors are solicited to contribute to the special issue by submitting articles that illustrate research results, projects, surveying works and industrial experiences that describe significant advances in the following areas, but are not limited to. Submissions may span a broad range of topics, e.g.:

Track A: Security

Access control, Anonymity, Audit and audit reduction & Authentication and authorization, Applied cryptography, Cryptanalysis, Digital Signatures, Biometric security, Boundary control devices, Certification and accreditation, Cross-layer design for security, Security & Network Management, Data and system integrity, Database security, Defensive information warfare, Denial of service protection, Intrusion Detection, Anti-malware, Distributed systems security, Electronic commerce, E-mail security, Spam, Phishing, E-mail fraud, Virus, worms, Trojan Protection, Grid security, Information hiding and watermarking & Information survivability, Insider threat protection, Integrity

Intellectual property protection, Internet/Intranet Security, Key management and key recovery, Language-based security, Mobile and wireless security, Mobile, Ad Hoc and Sensor Network Security, Monitoring and surveillance, Multimedia security ,Operating system security, Peer-to-peer security, Performance Evaluations of Protocols & Security Application, Privacy and data protection, Product evaluation criteria and compliance, Risk evaluation and security certification, Risk/vulnerability assessment, Security & Network Management, Security Models & protocols, Security threats & countermeasures (DDoS, MiM, Session Hijacking, Replay attack etc.), Trusted computing, Ubiquitous Computing Security, Virtualization security, VoIP security, Web 2.0 security, Submission Procedures, Active Defense Systems, Adaptive Defense Systems, Benchmark, Analysis and Evaluation of Security Systems, Distributed Access Control and Trust Management, Distributed Attack Systems and Mechanisms, Distributed Intrusion Detection/Prevention Systems, Denial-of-Service Attacks and Countermeasures, High Performance Security Systems, Identity Management and Authentication, Implementation, Deployment and Management of Security Systems, Intelligent Defense Systems, Internet and Network Forensics, Large-scale Attacks and Defense, RFID Security and Privacy, Security Architectures in Distributed Network Systems, Security for Critical Infrastructures, Security for P2P systems and Grid Systems, Security in E-Commerce, Security and Privacy in Wireless Networks, Secure Mobile Agents and Mobile Code, Security Protocols, Security Simulation and Tools, Security Theory and Tools, Standards and Assurance Methods, Trusted Computing, Viruses, Worms, and Other Malicious Code, World Wide Web Security, Novel and emerging secure architecture, Study of attack strategies, attack modeling, Case studies and analysis of actual attacks, Continuity of Operations during an attack, Key management, Trust management, Intrusion detection techniques, Intrusion response, alarm management, and correlation analysis, Study of tradeoffs between security and system performance, Intrusion tolerance systems, Secure protocols, Security in wireless networks (e.g. mesh networks, sensor networks, etc.), Cryptography and Secure Communications, Computer Forensics, Recovery and Healing, Security Visualization, Formal Methods in Security, Principles for Designing a Secure Computing System, Autonomic Security, Internet Security, Security in Health Care Systems, Security Solutions Using Reconfigurable Computing, Adaptive and Intelligent Defense Systems, Authentication and Access control, Denial of service attacks and countermeasures, Identity, Route and

Location Anonymity schemes, Intrusion detection and prevention techniques, Cryptography, encryption algorithms and Key management schemes, Secure routing schemes, Secure neighbor discovery and localization, Trust establishment and maintenance, Confidentiality and data integrity, Security architectures, deployments and solutions, Emerging threats to cloud-based services, Security model for new services, Cloud-aware web service security, Information hiding in Cloud Computing, Securing distributed data storage in cloud, Security, privacy and trust in mobile computing systems and applications, **Middleware security & Security features:** middleware software is an asset on its own and has to be protected, interaction between security-specific and other middleware features, e.g., context-awareness, **Middleware-level security monitoring and measurement:** metrics and mechanisms for quantification and evaluation of security enforced by the middleware, **Security co-design:** trade-off and co-design between application-based and middleware-based security, **Policy-based management:** innovative support for policy-based definition and enforcement of security concerns, **Identification and authentication mechanisms:** Means to capture application specific constraints in defining and enforcing access control rules, **Middleware-oriented security patterns:** identification of patterns for sound, reusable security, **Security in aspect-based middleware:** mechanisms for isolating and enforcing security aspects, **Security in agent-based platforms:** protection for mobile code and platforms, Smart Devices: Biometrics, National ID cards, Embedded Systems Security and TPMs, RFID Systems Security, Smart Card Security, Pervasive Systems: Digital Rights Management (DRM) in pervasive environments, Intrusion Detection and Information Filtering, Localization Systems Security (Tracking of People and Goods), Mobile Commerce Security, Privacy Enhancing Technologies, Security Protocols (for Identification and Authentication, Confidentiality and Privacy, and Integrity), Ubiquitous Networks: Ad Hoc Networks Security, Delay-Tolerant Network Security, Domestic Network Security, Peer-to-Peer Networks Security, Security Issues in Mobile and Ubiquitous Networks, Security of GSM/GPRS/UMTS Systems, Sensor Networks Security, Vehicular Network Security, Wireless Communication Security: Bluetooth, NFC, WiFi, WiMAX, WiMedia, others

This Track will emphasize the design, implementation, management and applications of computer communications, networks and services. Topics of mostly theoretical nature are also welcome, provided there is clear practical potential in applying the results of such work.

Track B: Computer Science

Broadband wireless technologies: LTE, WiMAX, WiRAN, HSDPA, HSUPA, Resource allocation and interference management, Quality of service and scheduling methods, Capacity planning and dimensioning, Cross-layer design and Physical layer based issue, Interworking architecture and interoperability, Relay assisted and cooperative communications, Location and provisioning and mobility management, Call admission and flow/congestion control, Performance optimization, Channel capacity modeling and analysis, Middleware Issues: Event-based, publish/subscribe, and message-oriented middleware, Reconfigurable, adaptable, and reflective middleware approaches, Middleware solutions for reliability, fault tolerance, and quality-of-service, Scalability of middleware, Context-aware middleware, Autonomic and self-managing middleware, Evaluation techniques for middleware solutions, Formal methods and tools for designing, verifying, and evaluating, middleware, Software engineering techniques for middleware, Service oriented middleware, Agent-based middleware, Security middleware, Network Applications: Network-based automation, Cloud applications, Ubiquitous and pervasive applications, Collaborative applications, RFID and sensor network applications, Mobile applications, Smart home applications, Infrastructure monitoring and control applications, Remote health monitoring, GPS and location-based applications, Networked vehicles applications, Alert applications, Embedded Computer System, Advanced Control Systems, and Intelligent Control : Advanced control and measurement, computer and microprocessor-based control, signal processing, estimation and identification techniques, application specific IC's, nonlinear and adaptive control, optimal and robot control, intelligent control, evolutionary computing, and intelligent systems, instrumentation subject to critical conditions, automotive, marine and aero-space control and all other control applications, Intelligent Control System, Wiring/Wireless Sensor, Signal Control System. Sensors, Actuators and Systems Integration : Intelligent sensors and actuators, multisensor fusion, sensor array and multi-channel processing, micro/nano technology, microsensors and microactuators, instrumentation electronics, MEMS and system integration, wireless sensor, Network Sensor, Hybrid

Sensor, Distributed Sensor Networks. Signal and Image Processing : Digital signal processing theory, methods, DSP implementation, speech processing, image and multidimensional signal processing, Image analysis and processing, Image and Multimedia applications, Real-time multimedia signal processing, Computer vision, Emerging signal processing areas, Remote Sensing, Signal processing in education. Industrial Informatics: Industrial applications of neural networks, fuzzy algorithms, Neuro-Fuzzy application, bioInformatics, real-time computer control, real-time information systems, human-machine interfaces, CAD/CAM/CAT/CIM, virtual reality, industrial communications, flexible manufacturing systems, industrial automated process, Data Storage Management, Harddisk control, Supply Chain Management, Logistics applications, Power plant automation, Drives automation. Information Technology, Management of Information System : Management information systems, Information Management, Nursing information management, Information System, Information Technology and their application, Data retrieval, Data Base Management, Decision analysis methods, Information processing, Operations research, E-Business, E-Commerce, E-Government, Computer Business, Security and risk management, Medical imaging, Biotechnology, Bio-Medicine, Computer-based information systems in health care, Changing Access to Patient Information, Healthcare Management Information Technology. Communication/Computer Network, Transportation Application : On-board diagnostics, Active safety systems, Communication systems, Wireless technology, Communication application, Navigation and Guidance, Vision-based applications, Speech interface, Sensor fusion, Networking theory and technologies, Transportation information, Autonomous vehicle, Vehicle application of affective computing, Advance Computing technology and their application : Broadband and intelligent networks, Data Mining, Data fusion, Computational intelligence, Information and data security, Information indexing and retrieval, Information processing, Information systems and applications, Internet applications and performances, Knowledge based systems, Knowledge management, Software Engineering, Decision making, Mobile networks and services, Network management and services, Neural Network, Fuzzy logics, Neuro-Fuzzy, Expert approaches, Innovation Technology and Management : Innovation and product development, Emerging advances in business and its applications, Creativity in Internet management and retailing, B2B and B2C management, Electronic transceiver device for Retail Marketing Industries, Facilities planning and management, Innovative pervasive computing applications, Programming paradigms for pervasive systems, Software evolution and maintenance in pervasive systems, Middleware services and agent technologies, Adaptive, autonomic and context-aware computing, Mobile/Wireless computing systems and services in pervasive computing, Energy-efficient and green pervasive computing, Communication architectures for pervasive computing, Ad hoc networks for pervasive communications, Pervasive opportunistic communications and applications, Enabling technologies for pervasive systems (e.g., wireless BAN, PAN), Positioning and tracking technologies, Sensors and RFID in pervasive systems, Multimodal sensing and context for pervasive applications, Pervasive sensing, perception and semantic interpretation, Smart devices and intelligent environments, Trust, security and privacy issues in pervasive systems, User interfaces and interaction models, Virtual immersive communications, Wearable computers, Standards and interfaces for pervasive computing environments, Social and economic models for pervasive systems, Active and Programmable Networks, Ad Hoc & Sensor Network, Congestion and/or Flow Control, Content Distribution, Grid Networking, High-speed Network Architectures, Internet Services and Applications, Optical Networks, Mobile and Wireless Networks, Network Modeling and Simulation, Multicast, Multimedia Communications, Network Control and Management, Network Protocols, Network Performance, Network Measurement, Peer to Peer and Overlay Networks, Quality of Service and Quality of Experience, Ubiquitous Networks, Crosscutting Themes – Internet Technologies, Infrastructure, Services and Applications; Open Source Tools, Open Models and Architectures; Security, Privacy and Trust; Navigation Systems, Location Based Services; Social Networks and Online Communities; ICT Convergence, Digital Economy and Digital Divide, Neural Networks, Pattern Recognition, Computer Vision, Advanced Computing Architectures and New Programming Models, Visualization and Virtual Reality as Applied to Computational Science, Computer Architecture and Embedded Systems, Technology in Education, Theoretical Computer Science, Computing Ethics, Computing Practices & Applications

Authors are invited to submit papers through e-mail ijcsiseditor@gmail.com. Submissions must be original and should not have been published previously or be under consideration for publication while being evaluated by IJCSIS. Before submission authors should carefully read over the journal's Author Guidelines, which are located at <http://sites.google.com/site/ijcsis/authors-notes> .

© IJCSIS PUBLICATION 2015

ISSN 1947 5500

<http://sites.google.com/site/ijcsis/>